



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
Main Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2019

---

## **Mischverteilung der Verlaufsmuster von Schulleistungen über die obligatorische Schulzeit**

Tomasik, Martin J ; Moser, Urs

**Abstract:** In diesem Aufsatz werden die Leistungszuwächse in Mathematik und Deutsch von Schülerinnen und Schülern während der obligatorischen Schulzeit betrachtet. Es wird untersucht, ob sich distinkte Typen von Leistungsverläufen identifizieren lassen, um wie viele Typen es sich handelt und ob sich die Zugehörigkeit zu einem Verlaufstyp durch Variablen auf der Ebene des Individuums, der Herkunftsfamilie und der Schulklasse vorhersagen lässt. Die Datengrundlage zur Untersuchung dieser Fragestellungen stellt die Zürcher Längsschnittstudie, in der die Schulleistungen von etwa 2000 Schülerinnen und Schülern drei, sechs und neun Jahre nach der Einschulung erfasst und mit IRT-Methoden auf einer gemeinsamen metrischen Skala abgebildet wurden. Die Auswertung erfolgte mit Hilfe von mischverteilten Wachstumskurvenmodellen, wobei fünf distinkte Typen von Schulleistungsverläufen identifiziert wurden, die als „durchschnittlicher Typ“ (20 %), „lernschwacher Typ“ (26 %), „Nachzüglertyp“ (9 %), „ertragsreicher Typ“ (30 %) sowie „Typ mit Vorsprung“ (15 %) bezeichnet werden können. Die verschiedenen Prädiktorvariablen waren in unterschiedlichem Ausmaß für die Vorhersage der Typenzugehörigkeit relevant, wobei die psychometrische Intelligenz, die vorschulischen Rechenfertigkeiten sowie selbständiges Arbeitsverhalten eine herausragende Rolle spielten.

DOI: <https://doi.org/10.1007/s11618-019-00895-z>

Other titles: A mixture distribution of school performance growth throughout compulsory schooling

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-179440>

Journal Article

Accepted Version

Originally published at:

Tomasik, Martin J; Moser, Urs (2019). Mischverteilung der Verlaufsmuster von Schulleistungen über die obligatorische Schulzeit. *Zeitschrift für Erziehungswissenschaft*, 22(4):811-849.

DOI: <https://doi.org/10.1007/s11618-019-00895-z>

Mischverteilung der Verlaufsmuster von Schulleistungen über die obligatorische Schulzeit  
[Mixture Distribution of School Performance Growth throughout Compulsory Schooling]

Martin J. Tomasik<sup>1,2</sup> und Urs Moser<sup>1</sup>

<sup>1</sup>Universität Zürich

<sup>2</sup>Universität Witten/Herdecke

Prof. Dr. Martin Tomasik, Institut für Bildungsevaluation, Assoziiertes Institut der Universität  
Zürich, Wilfriedstrasse 15, 8032 Zürich, Schweiz, martin.tomasik@ibe.uzh.ch

Prof. Dr. Urs Moser, Institut für Bildungsevaluation, Assoziiertes Institut der Universität  
Zürich, Wilfriedstrasse 15, 8032 Zürich, Schweiz, urs.moser@ibe.uzh.ch

### Zusammenfassung

In diesem Aufsatz werden die Leistungszuwächse in Mathematik und Deutsch von Schülerinnen und Schülern während der obligatorischen Schulzeit betrachtet. Es wird untersucht, ob sich distinkte Typen von Leistungsverläufen identifizieren lassen, um wie viele Typen es sich handelt und ob sich die Zugehörigkeit zu einem Verlaufstyp durch Variablen auf der Ebene des Individuums, der Herkunftsfamilie und der Schulklasse vorhersagen lässt. Die Datengrundlage zur Untersuchung dieser Fragestellungen stellt die Zürcher Längsschnittstudie, in der die Schulleistungen von etwa 2.000 Schülerinnen und Schülern drei, sechs und neun Jahre nach der Einschulung erfasst und mit IRT-Methoden auf einer gemeinsamen metrischen Skala abgebildet wurden. Die Auswertung erfolgte mit Hilfe von mischverteilten Wachstumskurvenmodellen, wobei fünf distinkte Typen von Schulleistungsverläufen identifiziert wurden, die als „durchschnittlicher Typ“ (20%), „lernschwacher Typ“ (26%), „Nachzüglertyp“ (9%), „ertragsreicher Typ“ (30%) sowie „Typ mit Vorsprung“ (15%) bezeichnet werden können. Die verschiedenen Prädiktorvariablen waren in unterschiedlichem Ausmaß für die Vorhersage der Typenzugehörigkeit relevant, wobei die psychometrische Intelligenz, die vorschulischen Rechenfertigkeiten sowie selbständiges Arbeitsverhalten eine herausragende Rolle spielten.

Schlüsselwörter: Deutsch; Geschlechterunterschiede; Verlaufstypen; Lernzuwachs; Mathematik; Schulleistungen

**Mixture Distribution of School Performance Growth throughout Compulsory Schooling**

We investigated the growth in school performance of students throughout compulsory schooling, examining whether distinct types of growth trajectories can be identified, how many there are, and whether the probability of type membership can be predicted by variables at the level of the individual, the family of origin, and the school class. Data come from the Zurich Learning Progress Study, which comprises a random sample of about 2000 students who have entered school in the 2003/04 school year. School performance in mathematics and language was assessed three, six, and nine years after school entry by standardized tests reflecting the official curriculum. These tests were scaled using probabilistic test theory and linked across the three measurement occasions so that they could be compared on a common metric scale covering most parts of compulsory schooling. Using non-linear growth mixture modelling, we identified five distinct types of performance growth, namely the “average type” (20%), the “slow learning type” (26%), the “left-behind type” (9%), the “profiting type” (30%), and the “head start type” (15%). The different predictor variables had different relevance for the different growth types, whereby psychometric intelligence, pre-schooling competencies in arithmetic, and independent learning and study habits played a prominent role.

Keywords: class composition effects; gender differences; language; learning trajectories; mathematics; school performance; type of learner

## 1 Einleitung

Schulleistungen sind häufig die Währung, in welcher der Bildungserfolg von einzelnen Schülerinnen und Schülern, aber auch ganzer Bildungsinstitutionen und Bildungssysteme gemessen wird. Dementsprechend viele Studien beschäftigen sich daher mit Ursachen von Schulleistungen (z. B. Gut, Reimann & Grob, 2012), betrachten damit zusammenhängende Prozesse und Korrelate (z. B. Martin & Marsh, 2006) oder untersuchen kurz- und langfristige Konsequenzen von Schulleistungen (z. B. Tomasik, Napolitano & Moser, in Druck). Viele dieser Studien – darunter ein überwiegender Teil nationaler und internationaler Vergleichsstudien – sind querschnittlich angelegt, in dem Sinne, dass Schulleistungen zu einem bestimmten Zeitpunkt gemessen werden und damit der aktuelle Lern- und Leistungsstand bestimmt wird. Obwohl dieser methodische Zugang forschungsökonomisch häufig notwendig ist (und manchmal auch den einzig machbaren darstellt) und obwohl diese Studien sehr viele relevante und spannende Einsichten ermöglicht haben, bieten sie immer nur eine Momentaufnahme und erlauben es nicht, die Frage zu beantworten, wie sich Schulleistungen über die Zeit entwickelt haben. Die Entwicklung über die Zeit aber kann, muss jedoch nicht, mit dem gegenwärtigen Leistungsstand korrespondieren. Dadurch verpassen querschnittlich angelegte Studien möglicherweise wichtige Variablen oder Prozesse, die mit Schulleistungen und deren Entwicklung in Zusammenhang stehen.

Längsschnittliche Studien eignen sich sehr viel besser für solche Fragestellungen und erlauben es durch die zeitliche Reihung der Daten außerdem noch, Prozesse in ihrem Verlauf beobachten zu können. Diese konzeptuellen und methodischen Vorteile müssen allerdings in der Regel mit einem großen organisatorischen Aufwand erkaufte werden, insbesondere wenn Schülerinnen und Schüler über längere Zeiträume beobachtet werden sollen, in denen Neuzusammensetzungen von Schulklassen, Schulwechsel und Umzüge stattfinden. Außerdem stellt sich bei längerfristigen Untersuchungen zwangsläufig die Frage nach der Invarianz von Messung und der Abbildung des sich in der Regel schnell verändernden Leistungsniveaus auf

einer vergleichbaren Skala. Aus diesem Grund verwundert es nicht, wenn die untersuchten Zeiträume der meisten Längsschnittstudien in der Regel kurz sind und selten mehr als einige wenige Schuljahre umfassen.

Aber selbst wenn eine Studie längsschnittlich angelegt ist und längere Zeiträume umfasst, ignoriert man in vielen Auswertungen die vorhandene Heterogenität in den Lernzuwachsen, indem von einem Durchschnittsverlauf ausgegangen wird, der dann für alle Schülerinnen und Schüler gelten soll. Etwaige Abweichungen vom Durchschnittsverlauf finden etwa in der Modellierung als Wachstumskurven zwar immer häufiger Verwendung, trotzdem wird aber auch hier von einem durchschnittlichen Ausgangsniveau und von einem durchschnittlichen Lernzuwachs für alle Schülerinnen und Schüler in einer Stichprobe ausgegangen. Auch kann die Nichtberücksichtigung einer zugrunde liegenden Typenstruktur unter Umständen zu falschen Schlussfolgerungen und Handlungsempfehlungen führen, wie ein zugegebenermaßen konstruiertes Beispiel in Abbildung 1 deutlich macht.

=====

ABBILDUNG 1 ETWA HIER

=====

Abhilfe schaffen hier Mischverteilungsmodelle, mit denen sich zunächst unterschiedliche Verlaufstypen identifizieren lassen und damit die Heterogenität in den Lernzuwachsen gebührend gewürdigt wird. Die so identifizierten Verlaufstypen werden dann in der Regel als qualitativ distinkte Gruppen behandelt, für die unterschiedliche Mechanismen und Zusammenhänge gelten können. Damit erlaubt der Mischverteilungsansatz nicht nur eine akkuratere Beschreibung der Lernzuwächse in heterogenen Schülerpopulationen, sondern kann auch hilfreich dafür sein, prototypische Verläufe abzubilden, die sich möglicherweise nicht ineinander überführen lassen (also dass das eine nicht lediglich ein mehr oder weniger des anderen ist), sondern qualitativ distinkt sind. Damit bieten solche Modelle die Möglichkeit, die erziehungswissenschaftliche oder pädagogisch-psychologische

Theoriebildung voranzutreiben.

Die vorliegende Studie verfolgt diesen Ansatz, indem sie verschiedene Typen der Verläufe von Schulleistungen in Mathematik und Deutsch unterscheidet und dafür Daten verwendet, welche die gesamte obligatorische Schulzeit umfassen. Das Ziel ist die empirische Identifikation, Beschreibung und Vorhersage der Typenzugehörigkeit. Sollten diese Typen auch in zukünftigen Untersuchungen so repliziert werden, wäre das ein erster Schritt in der Bestimmung längerfristiger „Lerntypen“ mit entsprechenden Konsequenzen für die Planung und Umsetzungen individualisierter pädagogischer Interventionen. Die verwendete Zufallsstichprobe erlaubt dabei eine gute Generalisierbarkeit der Befunde und garantiert eine gute Vergleichbarkeit mit zukünftigen Untersuchungen, in denen auch Zufallsstichproben Verwendung finden.

## **2 Stand der Forschung**

Bevor der vorliegende Ansatz systematisch vorgestellt wird, wollen wir zunächst den aktuellen Forschungsstand zu Lernverläufen über die gesamte Schulzeit und zu den wichtigsten Prädiktoren interindividueller Unterschiede in diesen Lernverläufen darstellen (für einen Überblick siehe auch Brühwiler, Helmke & Schrader, 2017). Um ein möglichst umfassendes Bild zu zeichnen, greifen wir dabei sowohl auf Studien aus dem deutschsprachigen Raum als auch auf international publizierte Studien zurück, wohl wissend, dass sich in Letzteren auch kulturelle und bildungspolitische Spezifika widerspiegeln, die sich im deutschsprachigen Raum so nicht finden.

### **2.1 Lernverläufe über die Schulzeit**

Zu Lernzuwachsen über die gesamte obligatorische Schulzeit liegen nur wenige Studien vor, von denen die meisten querschnittlich angelegt sind und somit streng genommen Unterschiede zwischen verschiedenen Jahrgangsstufen und nicht intraindividuelle Lernverläufe messen. Es gibt allerdings auch Studien, die als Längsschnitt oder zumindest als „beschleunigter Längsschnitt“ angelegt sind. Als Beispiel kann sicherlich die National

Education Longitudinal Study (NELS) genannt werden, in der mittlerweile zwei Kohorten sehr großer Stichproben von Schülerinnen und Schülern auf der Sekundarstufe und darüber hinaus untersucht wurden. Mit Daten der NELS konnten Muller, Stage und Kinzle (2001) beispielsweise zeigen, dass die Lernzuwächse auf der Sekundarstufe von verschiedenen Schülermerkmalen und vorherigen Leistungen abhängig waren.

Wei, Blackorby und Schiller (2011) verwenden das Design eines „beschleunigten Längsschnitts“, und zwar bei 7 bis 17 Jahre alten Schülerinnen und Schülern mit einer diagnostizierten Behinderung. Interessanterweise finden die Autoren für diese besondere Schülergruppe ein Abflachen des Lernzuwachses über die Zeit. An dieser Stelle interessanter erscheint allerdings, dass die Lernzuwächse systematisch als Funktion der Diagnose (etwa Lernbehinderung, Hörbeeinträchtigung usw.) unterscheiden.

Pfost, Hattie, Dörfler und Artelt (2014) haben in einer systematischen Übersicht Studien zusammengestellt, die interindividuelle Unterschiede bei intraindividuellen Lernzuwächsen im Lesen längsschnittlich untersucht haben. Alle diese Studien beschränken sich allerdings auf die Primarschulzeit und zeigen – entgegen der Erwartung der Autoren – keine Ausweitung der Leistungsunterschiede zwischen schwachen und starken Lesern. Beispielhaft sei die Studie von Geary (2011) erwähnt, in der der Lernzuwachs in Mathematik in Abhängigkeit von einer Reihe mathematischer Vorläuferfertigkeiten, der psychometrischen Intelligenz, dem Arbeitsgedächtnis und der kognitiven Verarbeitungsgeschwindigkeit untersucht wird.

Auch die Studie von Bloom, Hill, Black und Lipsey (2008) stammt aus dem nordamerikanischen Raum. Sie vergleicht sieben verschiedene Maße des Lernzuwachses im Lesen und umfasst vermutlich die breiteste Altersspanne von allen großen Studien, nämlich die vom Schuleintritt bis zum 12. Schuljahr. Es finden sich für auf 12 Monate angelegte Längsschnitte (auch Schuljahresäquivalente genannt) Lernzuwächse von  $-0.06 < d < 1.66$  Standardabweichungen mit einer deutlichen Abnahme vor allem in der zweiten Hälfte der



Schullaufbahn. Ähnliche Befunde finden sich übrigens für andere Kompetenzdomänen.

Allen bislang dargestellten Studien ist gemein, dass durchschnittliche Lernzuwächse für die Gesamtstichprobe ausgewertet und berichtet werden. Studien mit Mischverteilungen, in denen längere Zeiträume betrachtet werden und in Schulleistungen mit standardisierten Leistungstests erfasst werden, sind eher die Ausnahme. Wenn Mischverteilungen untersucht werden, dann beschränken sich die Zeiträume auf maximal einige Jahre und es werden Schulnoten oder andere, subjektive Leistungsmaße verwendet. Entsprechend finden sich im Ergebnis geringe bis keine Mittelwertveränderungen, sehr hohe Stabilitäten zwischen den einzelnen Messzeitpunkten und fast schon triviale Verlaufstypenkategorien. Morin, Rodriguez, Fallu, Maïano und Janosz (2011) finden beispielsweise bei Verwendung eines gewichteten Schulnotendurchschnitts und vier jährlichen Messzeitpunkten drei Gruppen: die konstant überdurchschnittlichen, die konstant durchschnittlichen und die konstant unterdurchschnittlichen Schülerinnen und Schüler.

Auch im deutschsprachigen Raum dominieren querschnittliche Lernstandserhebungen, und wenn längsschnittlich angelegte Studien vorliegen, dann umfassen sie nur kurze Zeiträume von einigen wenigen Schuljahren (z. B. Baumert, Gruehn, Heyn, Köller & Schnabel, 1997 [BIJU]; Baumert, Nagy & Lehmann, 2012 [ELEMENT]; Blossfeld, von Maurice & Schneider, 2011 [NEPS]; Pekrun et al., 2006 [PALMA]; Reiss, Klieme & Köller, 2017 [PISA Plus]). Zwei Ausnahmen bilden hier die Hamburger Studie zur Lernausgangslage und zur Lernentwicklung (LAU; Behörde für Schule und Berufsbildung, 2011, 2012), in der ein gesamter Jahrgang von Klassenstufe 5 bis 13 längsschnittlich auf die wichtigsten Aspekte seiner Lernentwicklung hin untersucht wurde, sowie das ebenfalls in Hamburg angesiedelte Kompetenzermittlungssystem (KERMIT; Lücken et al., 2014) mit längsschnittlichen Daten aus Jahrgangsstufen 2, 3, 5, 7, 8 und 9. Die einzige uns bekannte Studie im deutschsprachigen Raum, in der längsschnittliche Leistungsveränderungen über die *gesamte* Schulzeit hinweg untersucht wurden, stellt die Zürcher Lernstandserhebung (siehe Moser, Oostlander &

Tomasik, 2017) dar. In dieser Studie wurden die Schulleistungen an ein und derselben Stichprobe bei Einschulung sowie drei, sechs und neun Jahre danach gemessen und die letzten drei Messungen auf ein und derselben metrischen Skala abgebildet. Moser et al. berichten über Unterschiede in den Zuwächsen zwischen den beiden Schulfächern (Mathematik und Deutsch) sowie den unterschiedlichen Schultypen (Primarschule und Sekundarschule) und zeigen Prädiktoren für die Lernverläufe auf. Die implizite Annahme, die bei diesen Auswertungen gemacht wurde, war, dass es bei den Lernzuwächsen einen prototypischen Verlauf gäbe, der im Prinzip auf alle Schülerinnen und Schüler zutrifft und in dem sie sich nur in ihrem Ausgangsniveau unterscheiden. Unseres Wissens liegt für den deutschsprachigen Raum keine einzige Studie vor, die versucht hat, unterschiedliche Verlaufstypen zu identifizieren, zu beschreiben und durch Merkmale auf der Ebene der Schülerinnen und Schüler, ihrer Eltern und ihres Schulkontextes vorherzusagen.

Die vorliegende Untersuchung soll diese Forschungslücke schließen und greift dabei auf die Daten der Zürcher Lernstandserhebung zurück, die mit Hilfe von Mischverteilungsmodellen (z. B. Muthén & Asparouhov, 2007) untersucht werden. Der Einsatz solcher Modelle erlaubt es Verlaufstypen zu identifizieren, die sich voneinander unterscheiden, untereinander aber ähnlich sind. In einem ersten Schritt wird dabei die optimale Anzahl der Verlaufstypen aus den Daten heraus bestimmt (z. B. Nylund, Asparouhov & Muthén, 2007). In einem zweiten Schritt lässt sich dann die Wahrscheinlichkeit der Zugehörigkeit zu einem bestimmten Verlaufstyp mit Hilfe von Drittvariablen wie etwa dem Geschlecht oder dem Bildungshintergrund der Eltern bestimmen (z. B. Muthén & Muthén, 2000). Damit kombinieren Mischverteilungsmodelle personenorientierte mit variablenzentrierten Ansätzen und eignen sich dadurch besonders gut zur Identifikation interindividueller Unterschiede in intraindividuellen Veränderungen.

## **2.2 Interindividuelle Unterschiede in Lernverläufen**

Aus bisherigen querschnittlichen und längsschnittlichen Untersuchungen zu

Prädiktoren von Schulleistungen auf den verschiedenen Schuljahrgangsstufen sind verschiedene Drittvariablen bekannt, die mit Schulleistungen oder Lernzuwächsen korreliert sind. Im Folgenden findet sich eine Übersicht von Variablen, die auch wir für die Vorhersage der Verlaufstypen verwendet haben. Zum einen werden Variablen behandelt, die auf der Ebene der individuellen Schülerin oder des individuellen Schülers angesiedelt sind. Dabei handelt es sich um das Geschlecht, das Alter zum Zeitpunkt der Einschulung, die psychometrische Intelligenz, das schulbezogene Vorwissen sowie das Arbeits- und Sozialverhalten in der Schule. Zum anderen stellen wir Variablen vor, die sich auf den familiären Hintergrund der Schülerinnen und Schüler beziehen, nämlich die zu Hause gesprochene Sprache, den höchsten Bildungsabschluss der Eltern, das inner- und außerfamiliäre Freizeitverhalten sowie die Mediennutzung. Es wird nur in dem Umfang auf diese Variablen eingegangen, wie das für die Begründung der Hypothesen notwendig ist.

**Geschlecht.** Jungen und Mädchen unterscheiden sich systematisch in ihren Schulleistungen, aber diese Geschlechterdifferenzen sind kulturell verschieden und ändern sich über die Zeit hinweg. Während es in einigen Ländern gelingt, Geschlechterdifferenzen auszugleichen, ist dies beispielsweise in Deutschland noch nicht der Fall. Dabei zeigt sich in der Regel ein Vorteil der Jungen im mathematisch-naturwissenschaftlichen Bereich, während Mädchen im sprachlichen Bereich besser abschneiden (Budde, 2009). Jungen weisen dementsprechend in den internationalen Schulleistungsstudien, wie beispielsweise in PISA 2003, einen größeren Rückstand in Bezug auf die Lesekompetenz auf; in der mathematischen Kompetenz schneiden die Jungen dagegen etwas besser ab, die Unterschiede sind allerdings vergleichsweise klein (Kuhn, 2008, S. 54; Zimmer, Burba & Rost, 2004). Ein differenzierter Blick zeigt jedoch, dass Jungen vor allem über ein günstigeres Selbstkonzept verfügen und ihre Kompetenzen noch höher einschätzen als ihre Interessen. Kompetenzvorsprung, Selbstvertrauen und Interessen der Jungen wachsen bis zum Ende der Schule, so dass sich Mädchen seltener für mathematisch-technische Studiengänge entscheiden (Kuhn, 2008, S.

5f.). Allerdings lassen sich die generellen Geschlechterdifferenzen in der Sprachentwicklung und in der Mathematik nicht für alle Teilkompetenzen feststellen, was für andere Kompetenzbereiche, beispielsweise die Grob- und Feinmotorik, ebenfalls nachgewiesen werden konnte (Blossfeld, Bos, Hannover, Lenzen, Müller-Bölling, Prenzel & Wößmann, 2009, S. 65).

Ganz anders müssen die Geschlechterdifferenzen beurteilt werden, wenn die *schulischen Laufbahnen* von Jungen und Mädchen betrachtet werden. Hier haben Jungen mit größeren Verzögerungen zu kämpfen, werden also seltener vorzeitig eingeschult und häufiger von der Einschulung zurückgestellt (Konsortium Bildungsberichterstattung, 2006, S. 236). Außerdem müssen sie in manchen Jahrgangsstufen fast doppelt so häufig wie Mädchen eine Klassenstufe wiederholen (Krohne, Meier & Tillmann, 2004, S. 376 ff.) und verlassen das Bildungssystem später (Kuhn, 2008). Dieser Trend zu häufigeren Verzögerungen und Brüchen findet sich auch international und setzt sich in der Berufsbildung fort (OECD, 2015).

**Alter bei Einschulung.** Schulen sind in der Regel in Jahrgangsklassen organisiert. Mit einem Stichdatum wird dann festgelegt, zu welchem Jahrgang ein Kind gehört und wann seine Schullaufbahn beginnt. Der Stichtag beziehungsweise das Jahrgangsklassenprinzip hat den Effekt, dass Kinder, die kurz vor dem Stichtag Geburtstag haben, die Jüngsten in einer Klasse sind, während diejenigen, die kurz danach Geburtstag haben, ein weiteres Jahr auf den Schulstart warten (und sich dabei kognitiv, sozial und emotional weiterentwickeln). Bei der Einschulung im folgenden Jahr zählt es dann zu den älteren und häufig am weitesten entwickelten Kindern in der Klasse. Das relative Alter hängt mit verschiedenen Aspekten der Schullaufbahn zusammen. Bisherige Studien zeigen beispielsweise, dass das relative Lebensalter von Kindern und Jugendlichen mit der Leistungsentwicklung einhergeht. Bei gleichem Schuljahrgang erreichen jüngere Kinder in Schulleistungstests schlechtere Ergebnisse als ältere (z. B. Crawford, Dearden & Greaves, 2013). Das Alter hat jedoch nicht nur eine Bedeutung für die Schulleistungen, sondern auch für die Kontinuität der

Schullaufbahn. Jüngere Schüler erhalten schlechtere Noten (Cobley, McKenna, Baker & Wattie, 2009), mit geringerer Wahrscheinlichkeit eine Gymnasialempfehlung am Ende der Grundschule (Jürges & Schneider, 2006) und müssen eher eine Klasse wiederholen (Bedard & Dhuey, 2006; Puhani & Weber, 2007). Der Startvorteil älterer Schüler wird einerseits entwicklungsbiologisch erklärt. Ältere Kinder verfügen zu Beginn der Schulzeit über bessere kognitive, motivationale und emotionale Voraussetzungen für das schulische Lernen als jüngere Kinder, was sich in einer günstigeren Leistungsentwicklung auswirkt. Andererseits ermöglicht ein zusätzliches Jahr auch mehr Lernerfahrungen in der Familie, die zu einem Startvorteil führen und somit einen besseren Lernerfolg ermöglichen (vgl. Elder & Lubotsky, 2009; Gold, Duzy, Rauch & Quiroga Murcia, 2012; Lincove & Painter, 2006).

**Intelligenz.** Der Zusammenhang zwischen Intelligenz und Wissenserwerb ist sowohl konzeptuell als auch empirisch sehr eng (z. B. Gottfredson, 2002; Gustafsson & Undheim, 1996; Waldmann, Renkl & Gruber, 2003). So argumentiert beispielsweise Weinert (1996), dass „das Niveau der Intelligenz auch das kognitive Lernen beeinflusst, so dass sich mehr oder minder intelligente Lernprozesse ergeben. Deren Ertrag besteht im Erwerb eines intelligenteren oder weniger intelligenten Wissens“ (S. 96). Dass Intelligenz die Grundlage für jedes kognitive Lernen darstellt, lässt sich etwa sehr eindrücklich in Untersuchungen der Expertisenforschung (z. B. Schneider, Körkel & Weinert, 1989) oder auch in längsschnittlichen Analysen zur Entwicklung von Schulleistungen (z. B. Weinert & Schneider, 1999) zeigen. Dabei scheint nach Ackerman (1987, 1992) Intelligenz insbesondere zu Beginn des Kompetenzerwerbs relevant zu sein, wenn die kognitive Belastung noch hoch ist und noch keine Strategien zur Reduktion dieser Belastung entwickelt wurden.

Die Zusammenhänge zwischen Intelligenz und Schulleistungen sind von der Effektstärke her mittelhoch bis hoch. So finden beispielsweise Spinath, Freudenthaler und Neubauer (2010) bei Achtklässlern Korrelationen von über  $r = .50$ , wobei die Effekte für Mädchen etwas größer zu sein scheinen als für Jungen. Auch scheint zumindest in der

Primarschule der Effekt von Intelligenz auf die Mathematikleistungen stärker zu sein als der Effekt von Motivation, während für Deutschleistungen die Motivation stärkere Effekte hat (z. B. Weber, Lu, Shi & Spinath, 2013). Intelligenz könnte außerdem einen zentralen Vermittlungsmechanismus der weiter unten beschriebenen primären Herkunftseffekte darstellen (siehe Ferguson, Horwood & Boden, 2008), was für multivariate Auswertungen relevant ist.

**Schulbezogenes Vorwissen.** Kinder bringen von zu Hause, aus dem Kindergarten oder der Vorschule Wissen mit, das in engem Bezug zu den schulischen Anforderungen in der Schuleingangsphase und darüber hinaus steht. Die interindividuelle Varianz dieses Vorwissens ist dabei sehr groß. Während viele Kinder beispielsweise noch keine oder nur einzelne Buchstaben kennen, die sie noch nicht zu Wörtern verbinden können, können andere Kinder bereits einzelne Wörter oder sogar einfache Sätze lesen (z. B. Moser, Berweger & Stamm, 2005a). Eine ähnliche Heterogenität gibt es auch beim mathematischen Vorwissen (z. B. Moser, Berweger & Stamm, 2005b). Dabei zeigte sich in verschiedenen Bereichen, dass ein solches domänenspezifisches Vorwissen häufig ein noch stärkerer Prädiktor für Leistungen sein kann als Intelligenz (z. B. Ceci & Liker, 1986; Schneider & Björklund, 1992). Das gilt auch für den schulischen Bereich. So konnte beispielsweise Stern (1997) zeigen, dass die Mathematikleistungen auf der Sekundarstufe II enger mit den Mathematikleistungen in der Primarschule zusammenhing als mit der Intelligenz der Schülerinnen und Schüler. Zu ähnlichen Schlussfolgerungen kommen auch Watts, Duncan, Siegler und Davis-Keane (2014), die mathematisches Vorwissen im Alter von nur 54 Monaten untersucht haben und zeigen konnten, dass dieses Vorwissen Mathematikleistungen im Alter von 15 Jahren selbst nach Kontrolle von frühen Lesekompetenzen, Intelligenz sowie einer Reihe von Eltern- und Schülermerkmalen vorhersagen konnte. Das schulbezogene Vorwissen scheint damit ein starker Prädiktor für langfristige Lernverläufe zu sein.

**Arbeits- und Sozialverhalten.** Im Rahmen eines Angebot-Nutzungs-Modells (z. B.

Helmke, 2015) zählt das Arbeits- und Sozialverhalten neben der Intelligenz und dem Vorwissen zu den wichtigsten individuellen Bedingungsfaktoren des Lernerfolgs. Hier spiegeln sich, je nach Definition, metakognitive Fähigkeiten, Lernstile, die Steuerung der Aufmerksamkeit oder auch die Fähigkeit zu kooperativem Lernen wider (vgl. Brühwiler et al., 2017). Vor allem wenn das Arbeits- und Sozialverhalten nicht über Selbstberichte, sondern über eine konkrete Aufgabe oder über Fremdbeobachtungen erfasst wird, zeigen sich konsistente Zusammenhänge mit Lernerfolg (z. B. Abt Gürber, Bucchini & Brühwiler, 2011; Klauer & Leutner, 2007; Matthews, Zeidner & Roberts, 2006). Dies gilt insbesondere für die Fähigkeit zum selbständigen, selbstregulierten Lernen. So haben eine Reihe von Autoren positive Zusammenhänge zwischen Selbstregulation und Schulleistungen nachgewiesen, und zwar nicht nur korrelativ (z. B. Chung, 2000; Paris & Paris, 2001; Winnie, 1995; Zimmermann, 1990), sondern auch in einer Reihe von sorgfältig geplanten und durchgeführten Interventionsstudien (siehe dazu Dignath, Buettner & Langfeldt, 2008; Donker, de Boer, Kostons, Dignath van Ewijk & van der Werf, 2014). Nicht weniger bedeutsam sind Aspekte des Sozialverhaltens, und zwar sowohl im Hinblick auf antisoziale (z. B. Hinshaw, 1992) und prosoziale Verhaltensweisen (z. B. Wentzel, 1993; Caprara, Barbanelli, Pastorelli, Bandura & Zimbardo, 2000). Hierbei scheinen die unterschiedliche Beliebtheit der Schülerinnen und Schüler in der Klasse (Becherer, Köller & Zimmermann, 2017) und arbeitsvermeidendes Verhalten dissozialer Schülerinnen und Schüler (Metsäpelto, Pakarinen, Kiuru, Poikkeus, Lerkkanen & Nurmi, 2015) eine vermittelnde Rolle zwischen Sozialverhalten und Schulleistungen zu spielen.

**Erstsprache.** Das Beherrschen der Unterrichtssprache stellt eine der wichtigsten Lernvoraussetzungen dar, um sich am Unterricht beteiligen zu können und um letztlich vom Unterricht zu profitieren (Zöllner, Roos & Schöler, 2006). Als theoretische Erklärung für den in einer Vielzahl von Untersuchungen nachgewiesenen Nachteil von Kindern, deren Erstsprache nicht die Unterrichtssprache ist (z. B. Baumert & Schümer, 2001; Bos et al.,

2003; Schnabel & Schwippert, 2000; Schröder-Lenzen & Merrens, 2006), wird häufig die These der Entwicklungsinterdependenz von Cummins (1991) ins Spiel gebracht. Hierbei geht man davon aus, dass sprachliche und kognitive Entwicklung eng zusammenhängen, so dass die kognitive Entwicklung gefährdet ist, wenn die bilingualen Schülerinnen und Schüler es nicht schaffen, sowohl in der Herkunfts- als auch in der Zweitsprache einen ausreichend elaborierten Sprachstand zu erreichen. Hinzu kommt, dass das Erlernen einer Zweitsprache zunächst kognitive Ressourcen kostet und zusätzliche Lernzeit erfordert, die an anderer Stelle fehlt (Hopf, 2005). Eine weitere, ergänzende Erklärung für eine nachteilige schulische Entwicklung von fremdsprachigen Kindern gibt die Humankapitalhypothese. Sie beruht auf der Beobachtung, dass Migration (und damit Fremdsprachigkeit) häufig mit einem durchschnittlich niedrigeren sozioökonomischen Status konfundiert ist (z. B. Ramm, Prenzel, Heidemeier & Walter, 2004), der ursächlich für die nachteilige Entwicklung der Schulleistungen ist. Für diese Hypothese würde sprechen, dass der Unterschied in den Leistungen zwischen einheimischen und zugewanderten Kindern in den Ländern relativ niedrig ausfällt, die eine restriktive und hochselektive Einwanderungspolitik ausüben (Ramm et al., 2004, S. 260). Diese Beobachtung führt uns zur Frage nach der Rolle der sozialen Herkunft für den Bildungserfolg und den damit verbundenen Mechanismen.

**Soziale Herkunft.** Der Zusammenhang zwischen sozialer Herkunft und Schulleistungen wird vor allem im Rahmen der internationalen Schulleistungsvergleiche regelmäßig nachgewiesen (OECD, 2016). Zur Erklärung, weshalb und in welcher Form die soziale Herkunft für den Schulverlauf und den unterschiedlichen Kompetenzerwerb ist, ist Boudons Unterscheidung in primäre und sekundäre Herkunftseffekte grundlegend (Boudon, 1974; Maaz, Baumert & Trautwein, 2009, S. 14). Als primäre Herkunftseffekte werden Effekte der sozialen Herkunft bezeichnet, die sich direkt auf die Kompetenzentwicklung der Kinder auswirken und sich in den Schulleistungen zeigen. Dementsprechend sind unterschiedliche Entwicklungen im Kompetenzerwerb die Folge der unterschiedlichen



Ausstattung von Familien mit ökonomischem, sozialem und kulturellem Kapital. Primäre Herkunftseffekte sind Effekte, die sich aufgrund der (fehlenden) Unterstützung durch die Eltern im Laufe der Schulzeit in den Schulleistungen manifestieren. Sie werden als durch die Leistungen gedeckte soziale Ungleichheiten bezeichnet (Bourdieu, 1973; Bourdieu & Passeron, 1971; Maaz et al., 2009). Sekundäre Herkunftseffekte zeigen sich dagegen im Entscheidungsverhalten von Eltern an den Schnittstellen des Bildungssystems. Sekundäre Ungleichheiten sind soziale Ungleichheiten, die – bei gleichen schulischen Leistungen – aufgrund von je nach sozialer Herkunft der Familie spezifischen Übergangsentscheidungen während der Schullaufbahn zustande kommen (Boudon, 1974; Maaz et al., 2010).

Für den Kompetenzerwerb sind die primären Herkunftseffekte entscheidend. Es ist davon auszugehen, dass Schülerinnen und Schüler mit privilegierter sozialer Herkunft mehr Unterstützung für den Kompetenzerwerb erhalten als solche mit benachteiligender sozialer Herkunft. Im Vergleich zu anderen Herkunftsmerkmalen wie Migrationshintergrund und Erstsprache hat die soziale Herkunft der Schülerinnen und Schüler einen stärkeren Einfluss auf die Leistungsentwicklung. Schülerinnen und Schüler mit benachteiligender sozialer Herkunft starten nicht nur auf einem niedrigeren Ausgangsniveau; es zeigt sich bei der Leistungsentwicklung auch ein Schereneffekt. Zudem lassen sich bei der Schnittstelle zwischen Primarschule und Sekundarstufe I zusätzlich sekundäre Herkunftseffekte nachweisen. Bei gleichen Leistungen zeigt sich die soziale Herkunft hochgradig prädiktiv für den Übertritt in eine Schule mit höheren Anforderungen (z. B. Gymnasium; siehe Moser et al., 2017; Maaz et al., 2009).

**Förderung in der Familie.** Sozialisatorische Aktivitäten der Eltern gewinnen in den ersten Lebensjahren in den meisten Familien eine zunehmend große Bedeutung, womit der häusliche Anregungsgehalt immer bedeutsamer wird (z. B. Bradley & Corwyn, 2006). Breit verstanden umfasst der häusliche Anregungsgehalt sowohl materielle als auch kulturelle Aspekte wie elterliche Responsivität, Vielfalt der Stimulation, Auswahl und Qualität von

Spielzeug und Lernmaterial oder das Ausmaß und die Qualität der sprachlichen Interaktion (vgl. Totsika & Sylva, 2004). Zweifelsohne korrelieren diese Merkmale der Förderung in der Familie mit der kognitiven, sprachlichen und sozialen Entwicklung beim Schuleintritt (z. B. NICHD Early Child Care Research Network, 2005) und beeinflussen so auch das spätere schulische Lernen (z. B. Baker & Scher, 2002; Gottfried, Fleming & Gottfried, 2001; McElvany, Becker & Lüdtke, 2009).

**Freizeitaktivitäten.** Freizeitaktivitäten von Kindern lassen sich in zwei Kategorien unterscheiden, nämlich strukturierte und unstrukturierte Aktivitäten. Strukturierte Aktivitäten zeichnen sich dadurch aus, dass sie eine Regelmäßigkeit in der Teilnahme erfordern oder zumindest einfordern, auf gewissen Regeln aufbauen, durch einen oder mehrere erwachsene Aufsichtspersonen angeleitet oder zu mindestens beaufsichtigt werden, auf den Aufbau von Kompetenzen ausgerichtet sind, die in Schwierigkeitsgrad und Komplexität zunehmen, handlungsbasiert sind und damit eine aktive Mitarbeit erfordern sowie, schließlich, eine klare Rückmeldung zur Leistung erlauben (z. B. Mahoney, 2000). In Sportvereinen, bei den Pfadfindern oder im Kinderchor beispielsweise sind diese Merkmale alle gegeben. Bei unstrukturierten Aktivitäten, wie etwa Musik hören, mit Freunden zusammen sein, aber auch für sich selber Lesen fehlen diese Merkmale ganz oder teilweise. Die meiste Forschung zum Zusammenhang zwischen dem Freizeitverhalten und psychosozialer Anpassung gibt es für das Jugendalter. Dort findet man immer wieder, dass unstrukturierte Aktivitäten mit antisozialem Problemverhalten, Kontakt zu devianten Peers oder schlechten Eltern-Kind-Beziehungen einhergehen, während strukturierte Freizeitaktivitäten mit höheren Schulleistungen und insgesamt positiveren Entwicklungsergebnissen (wie etwa interpersonelle Kompetenz, höheres Selbstkonzept, höhere Überzeugungen bezüglich der eigenen Selbstwirksamkeit, Schulengagement oder Bildungsaspirationen) einhergehen (z. B. Eccles & Barber, 1999; Eccles & Templeton, 2002; Mahoney, Cairns & Farmer, 2003; Mahoney & Stattin, 2000). Vereinzelt finden sich allerdings auch Studien, die in

Freizeitaktivitäten an sich attraktive Ablenkungen sehen, die durch Zeit- und Aufmerksamkeitskonflikte mit schlechteren Lernergebnissen in Zusammenhang stehen (z. B. Marsh, 1992). Zudem ist die Frage nach der Wirkungsrichtung und der Selbstselektion nicht endgültig geklärt (z. B. Duncan, Magnuson & Ludwig, 2004).

**Mediennutzung.** Zur Rolle der Mediennutzung für die soziale, emotionale und kognitive Entwicklung von Kindern und Jugendlichen gibt es eine lang anhaltende Debatte über nachteilige Effekte insbesondere auf die Sprach- und Lesefähigkeit (z. B. Ennemoser, Schiffer, Reinsch & Schneider, 2003), die an dieser Stelle nicht in ihrer Gänze dargestellt werden kann. Zum einen wird argumentiert, dass Mediennutzung Zeit koste, die für andere Freizeitaktivitäten fehlen würde, dass die genutzten Medieninhalte selbst nachteilig, da häufig altersunangemessen gewalttätig seien oder dass der Medienkonsum den Transfer von schulisch gelerntem Wissen ins Langzeitgedächtnis erschwere (siehe Mößle, Kleimann & Rehbein, 2007). Andererseits ist der kausale Zusammenhang zwischen Mediennutzung und Schulleistungen alles andere als klar (z. B. Gentile, Swing, Lim & Khoo, 2012) und es wird in die Waage geworfen, dass Medien auch Gelegenheiten für informelles Lernen bieten würden, die sich wiederum positiv auf die Schulleistungen auswirkten (z. B. Buckingham & Willett, 2013; Tully, 1994). Allerdings kann man auch hier argumentieren, dass die Auswahl der Medieninhalte und deren Qualität für informelles Lernen sozial selektiv ist und sich dadurch soziale Ungleichheiten in Bezug auf Bildung reproduzieren (vgl. Stecher, 2005).

### 3 Fragestellung

Die vorliegende Untersuchung wird von vier Forschungsfragen geleitet. (1) Lassen sich distinkte Verlaufstypen von Lernzuwächsen über die gesamte obligatorische Schulzeit identifizieren und, wenn ja, wie viele Verlaufstypen gibt es? (2) Wie lassen sich die unterschiedlichen Verlaufstypen beschreiben? (3) Wie stellt sich die anteilmäßige Verteilung der Schülerinnen und Schüler an den unterschiedlichen Verlaufstypen dar? Und (4) lässt sich die Zugehörigkeit zu den Verlaufstypen durch unterschiedliche Variablen bzw.

Variablenkombinationen vorhersagen und, wenn ja, welche Variablen bzw.

Variablenkombinationen sind das? Bei der Beantwortung der letzten Frage sollen Variablen überprüft werden, für die es aus der bisherigen, oben skizzierten Forschung Hinweise dafür gibt, dass sie mit individuellen Lernverläufen – was abzugrenzen ist von distinkten Lernverlaufstypen – in Verbindung stehen.

## **4 Methode**

### **4.1 Studie und Stichprobe**

Die Grundgesamtheit der Längsschnittstichprobe der Zürcher Lernstandserhebung bildeten alle 11.118 Schülerinnen und Schüler, die im Sommer 2003 im Kanton Zürich in eine der 650 1. Klassen eingetreten sind. Ausgenommen waren davon 1. Klassen in Sonderschulen, Reformschulen sowie Mehrklassenschulen. Da die Lernstandserhebung nicht nur zum Ziel hatte, eine repräsentative Beschreibung des Lern- und Entwicklungsstandes im Verlauf der obligatorischen Schulzeit zu ermöglichen, sondern auch die Wirkungen von zwei Reformprojekten zu evaluieren sowie den Verlauf des Lernfortschritts in sogenannten Kleinklassen zu untersuchen, wurde die Stichprobe stratifiziert. Entsprechend wurden vier Subpopulationen definiert und bei der Stichprobenbildung berücksichtigt.

Die zahlenmäßig stärkste Gruppe bestand aus Schülerinnen und Schülern, die im Schuljahr 2003/04 in die 1. Regelklasse eintraten, wobei aus organisatorischen Gründen nur jene Kinder der Population angehörten, die nicht in eine Mehrklassenschule eintraten. Die zweite und dritte Gruppe bestand aus Schülerinnen und Schülern, die im gleichen Schuljahr in die 1. Regelklasse des Reformprojekts „Teilautonome Volksschule“ (vgl. Rhyn, Widmer, Roos & Nideröst, 2002) bzw. einer Schule eintraten, die am Reformprojekt „Qualität in multikulturellen Schulen“ (vgl. Bildungsdirektion Kanton Zürich, 2008) teilnahm. Die vierte Gruppe bestand aus Schülerinnen und Schülern, die im Schuljahr 2003/04 in die 1. Klasse der Kleinklasse A (vgl. Bildungsdirektion Kanton Zürich, 2010) eintraten. In der Kleinklasse A wird der Unterrichtsstoff der 1. Klasse auf zwei Schuljahre verteilt und Kinder, die am Ende

des zweiten Schuljahres die Lernziele der 1. Regelklasse erreichen, werden anschließend in die 2. Regelklasse versetzt.

Die Stichproben für die 1. Regelklassen sowie für die Kleinklassen A wurden nach einem zweistufigen Auswahlverfahren gebildet. Im ersten Schritt wurden die Klassen aufgrund des Sozialindex der jeweiligen Schulgemeinde und der Klassenstärke geordnet, wobei beide Informationen für eine implizite Stratifizierung (vgl. Satin & Shastry, 1990, S. 20) verwendet wurden. Im zweiten Schritt wurde die Klassen proportional zur Anzahl der Schülerinnen und Schülern, die zu Beginn des Schuljahres den Klassen zugeteilt waren, ausgewählt. Wie aus Tabelle 1 ersichtlich wird, wurden dabei die Klassen aus den beiden Reformprojekten sowie die Kleinklassen überrepräsentiert. Von den ausgewählten Klassen wurden anschließend sämtliche 2.043 Schülerinnen und Schüler in die Lernstandserhebung aufgenommen und zur Teilnahme eingeladen. Davon nahmen effektiv  $N_1 = 1970$  Schülerinnen und Schüler an der Erhebung zum Anfang des Schuljahres 2003/04 teil, was einer Teilnahmequote von 96.3% entspricht. Weitere Details zum Stichprobenverfahren finden sich bei Moser und Stamm (2005).

=====

TABELLE 1 ETWA HIER

=====

Am Ende des Schuljahres 2005/06 (T2) wurde die Stichprobe aktualisiert und reduzierte sich vor allem durch Wegzug von Schülerinnen und Schülern aus dem Kanton Zürich auf nunmehr 1.960 Kinder, die alle zur Teilnahme an der zweiten Lernstandserhebung eingeladen wurden und zwar unabhängig davon, ob sie noch die gleiche Schulklasse oder Schule besuchten oder nicht. Wie auch zu allen folgenden Messzeitpunkten wurden alle Kinder individuell nachverfolgt, vorausgesetzt, sie wohnten immer noch im Kanton. Von diesen Kindern nahmen schlussendlich  $N_2 = 1935$  effektiv teil. Weitere Details zur Stichprobe zu diesem Messzeitpunkt finden sich bei Keller und Moser (2008).

Am Ende des Schuljahres 2008/09 (T3) hatte sich die Stichprobe durch Wegzug von weiteren 157 Kindern aus dem Kanton Zürich auf 1803 Schülerinnen und Schüler reduziert, von denen insgesamt  $N_3 = 1679$  Schülerinnen und Schüler an der dritten Lernstandserhebung teilnahmen. Weitere Details zur Stichprobe zum Messzeitpunkt T3 finden sich bei Angelone und Moser (2011a).

Die vierte und letzte Lernstandserhebung (T4) fand am Ende des Schuljahres 2011/12 statt und basierte auf einer Stichprobe von 1817 Schülerinnen und Schülern. Zu diesem Messzeitpunkt waren diese auf unterschiedliche Schulen, Schultypen, Klassen und Klassenstufen aufgeteilt, zum einen wegen des erfolgten Übertritts auf eine weiterführende Schule und zum anderen wegen eines möglichen Wiederholens oder Überspringens einer Klassenstufe. Die Verteilung der Jugendlichen auf die unterschiedlichen Schultypen und Klassenstufen findet sich in Tabelle 2, woraus ersichtlich wird, dass die große Mehrheit von 81.0% der Schülerinnen und Schüler sich zu diesem Messzeitpunkt in der 9. und damit letzten obligatorischen Klassenstufe befand. Weitere 17.9% wiederholten eine Klassenstufe und 1.1% waren bereits in der 10. Klasse. An der vierten Erhebung nahmen insgesamt  $N_4 = 1635$  Schülerinnen und Schüler teil. Weitere Details zur Stichprobe zum vorläufig letzten Messzeitpunkt finden sich bei Angelone, Keller und Moser (2013, Kap. 2.1).

Letztlich gab es von  $N_{\text{brutto}} = 1896$  Schülerinnen und Schülern zu zumindest einem Messzeitpunkt verwendbare Daten zu deren Schulleistungen. Effektiv verwendet wurden für die Analysen Daten von  $N_{\text{netto}} = 1769$  Schülerinnen und Schülern nach Ausschluss multivariater Ausreißer (Mahalanobis  $p < .05$ ), die auch deswegen ausgeschlossen wurden, weil sie eine stabile Schätzung der Mischverteilungsmodelle unmöglich machten. Bei allen hier berichteten Auswertungen und Ergebnissen (mit Ausnahme von Tabelle 1) verwenden wir gewichtete Daten, bei denen die bei der Stichprobenziehung erfolgte Stratifizierung (vgl. Tabelle 1) statistisch korrigiert wird und somit Populationsparameter geschätzt werden.

=====

TABELLE 2 ETWA HIER

=====

## 4.2 Operationalisierung der Schulleistungen

**Aufgaben.** Für die Bestimmung des Lernstandes in der 3. Klasse (T2) wurden auf Grundlage des Lehrplans und der Lehrmittel des Kantons Zürich jeweils ein Deutschtest und ein Mathematiktest durch qualifizierte Lehrpersonen entwickelt und, wie alle folgenden Leistungstests auch, durch das Institut für Bildungsevaluation erprobt. Der Deutschtest umfasst die vier Lehrplanbereiche „Texte lesen und verstehen“, „Sprachbetrachtung“, „Texte überarbeiten“ und „Texte für sich und andere schreiben“. Es wurden insgesamt 115 Items mit einer mittleren Schwierigkeit von  $P = .55$  und einer mittleren Trennschärfe von  $r_{it} = .49$  ausgewertet. Die (immer jeweils unkonditioniert berechnete) EAP-Reliabilität lag bei  $\rho = .90$ . Der Mathematiktest umfasste die Lehrplanbereiche „Größen“, „Operationen“, „Mengen“, „Zahlenbereich“, „Zahlenschreibweise“ sowie „Geometrie“. Es gab insgesamt 167 Items mit einer mittleren Schwierigkeit von  $P = .62$  und einer mittleren Trennschärfe von  $r_{it} = .56$ . Die EAP-Reliabilität lag bei  $\rho = .93$ . Eine Auswahl von Aufgaben- und Lösungsbeispielen findet sich bei Keller und Moser (2008a, 2008b).

In der 6. Klasse (T3) umfasste der Deutschtest die Lehrplanbereiche „Leseverständnis“, „Wortschatz“, „Sprachbetrachtung“ und „Texte überarbeiten“ und bestand aus 326 Items mit einer mittleren Schwierigkeit von  $P = .60$  und einer mittleren Trennschärfe von  $r_{it} = .36$ . Die EAP-Reliabilität lag bei  $\rho = .95$ . Im Mathematiktest wurden die Lehrplanbereiche „Arithmetik“, „Größen/Sachrechnen/Proportionen“, „Brüche“, „Problemlösen“ und „Geometrie“ erfasst. Es wurden insgesamt 179 Items mit einer mittleren Schwierigkeit von  $P = .59$  und einer mittleren Trennschärfe von  $r_{it} = .44$  ausgewertet. Die EAP-Reliabilität lag bei  $\rho = .88$ . Eine Auswahl von Aufgabenbeispielen findet sich bei Moser und Angelone (2011).

Auch am Ende der obligatorischen Schulzeit (T4) wurden die Schulleistungen in den

Kernfächern Deutsch und Mathematik erfasst, allerdings wurden für die verschiedenen Schultypen drei unterschiedliche Testheftversionen mit unterschiedlicher Aufgabenschwierigkeit eingesetzt (Multi-Matrix-Design, vgl. Gonzales & Rutkowski, 2010). Mit dem Leistungstest in Deutsch wurden die Fähigkeiten der Schülerinnen und Schüler in den Lehrplanbereichen „Lesen: Texte lesen und verstehen“ und „Texte überarbeiten: Sprachbetrachtung, Grammatik, Rechtschreibung“ geprüft. Es wurden insgesamt 184 Items mit einer mittleren Schwierigkeit von  $P = .62$  und einer mittleren Trennschärfe von  $r_{it} = .32$  ausgewertet. Die EAP-Reliabilität lag bei  $\rho = .95$ . In Mathematik wurden die Lehrplanbereiche „Operation/Algebra“, „Gleichungen/Ungleichungen“, „Größen/Sachrechnen“, „Proportionalität/Funktionen“ sowie „Geometrie“ abgedeckt. Bei 85 Items betrug die mittlere Schwierigkeit  $P = .37$  und die mittlere Trennschärfe  $r_{it} = .38$ . Die EAP-Reliabilität lag bei  $\rho = .87$ . Eine Auswahl von Aufgabenbeispielen findet sich bei Angelone et al. (2013, Kap. 2.2.2).

**Skalierung.** Die Leistungstests wurden nach dem probabilistischen Raschmodell (siehe Bond & Fox, 2015) skaliert und auf dieser Grundlage wurden die Fähigkeiten nach der WLE-Methode (siehe Warm, 1989) geschätzt. Die Passung der Items zu dem Modell wurde mit Hilfe der *weighted mean square fit statistic* (siehe Wright & Masters, 1982, S. 99) bestimmt. In Ergänzung dazu wurden Items mit extremen Schwierigkeiten sowie Items mit einer niedrigen Trennschärfe (Item-Test-Korrelation) ausgeschlossen. Außerdem wurden die Itemcharakteristiken visuell untersucht und nur solche Items beibehalten, die über das gesamte Fähigkeitskontinuum unauffällig waren.

Das *linking* der Leistungstests über das 3., 6. und 9. Schuljahr auf eine gemeinsame metrische Skala erfolgte mit Hilfe des *common item non-equivalent groups design* (siehe Kolen & Brennan, 2004). Weil der Lernfortschritt in den dazwischen liegenden jeweils drei Schuljahren zu groß für ein direktes *linking* war, wurden angepasste Leistungstests zusätzlichen Stichproben aus dazwischen liegenden Klassenstufen (d.h. aus der 4., 5., 7. und



8. Klassenstufe) vorgelegt. Die Kalibrierungstichproben umfassten jeweils etwa 150 bis 180 Schülerinnen und Schüler pro Klassenstufe und die Zahl der verwendeten *linking*-Items belief sich auf 26 bis 215 in Abhängigkeit von Schulfach und Klassenstufe. Für jede dazwischen liegende Klassenstufe (z. B. Klassenstufe 4) wurde eine eigene Kalibrierungstichprobe gezogen, der dann ein Test vorgelegt wurde, der neben eigenen Items auch Linkitems zu darunter (z. B. Klassenstufe 3) und darüber (z. B. Klassenstufe 5) liegenden Klassenstufen beinhaltete. Dabei wurden nach Möglichkeit immer alle Lehrplanbereiche abgedeckt, die auch in der Hauptuntersuchung getestet wurden.

Die Leistungstests im 9. Schuljahr wurden in drei unterschiedlichen Versionen angeboten, da die Bildungsangebote in Zürich ab dem 7. Schuljahr dreigegliedert sind (Sekundarschule B/C, Sekundarschule A und Gymnasium). Die Items aus den drei Versionen wurden auf *differential item functioning* mit Hilfe eines Mehrfacetten-Raschmodells (siehe Lincare, 1994) getestet. Es wurden im Folgenden nur solche Items verwendet, bei der kein signifikanter Item  $\times$  Schultyp-Interaktionseffekt vorlag. Der Einfluss der drei unterschiedlichen Lehrpläne auf die Testscores wurde auch dadurch minimiert, dass nur Themen geprüft wurden, die in allen drei Schultypen behandelt worden sind. Alles zusammengekommen wurden die die Leistungstest so skaliert, dass ein Vergleich über die Zeit und über die verschiedenen Schultypen auf einem intervallskalierten Niveau möglich war.

Die deskriptiven Befunde zu den einzelnen WLEs sowie zu deren Korrelationen untereinander finden sich in Tabelle 3. Detailliertere Informationen können der zitierten Literatur entnommen werden.

---

---

TABELLE 3 ETWA HIER

---

---

**Modellierung.** Ausgangspunkt der Modellierung war ein latentes

Wachstumskurvenmodell zweiter Ordnung, das für beide Schulfächer simultan geschätzt wurde und das in Abbildung 2 dargestellt ist. Die Modellierung erfolgte in Mplus 6.12 (Muthén & Muthén, 2010) mit Hilfe des robusten Maximum Likelihood-Schätzers (MLR) unter Berücksichtigung der Stratifizierung mit Hilfe des STRATIFICATION-Befehls. Die Daten wurden mit einem normalisierten Gewicht auf Individualebene gewichtet, um die unterschiedliche Ziehungswahrscheinlichkeit der verschiedenen Schultypen (siehe Tabelle 1) zu berücksichtigen. Weil ganze Schulklassen gezogen wurden, korrigierten wir die Standardfehler mit Hilfe des COMPLEX-Befehls, wobei die Zugehörigkeit zu einer der 120 Schulklassen die CLUSTER-Variable darstellte. Fehlende Werte wurden durch FIML berücksichtigt.

=====

ABBILDUNG 2 ETWA HIER

=====

Um der Möglichkeit einer nichtlinearen Entwicklung der Schulleistungen Rechnung zu tragen (siehe Bloom et al., 2008) wurde die jeweils dritte Ladung  $\lambda_9$  des schulfachspezifischen Wachstumsfaktors aus den Daten frei geschätzt. Dieses nicht-lineare Modell zeigte eine sehr gute Passung zu den Daten ( $\chi^2 = 51.92$ ,  $df = 10$ ,  $p < .001$ ; RMSEA = .049; CFI = .996; TLI = .994; SRMR = .085) und wurde für die folgenden Auswertungen verwendet, zumal die Passung eines linearen Modells deutlich schlechter war ( $\Delta\chi^2 = 1538.81$ ,  $df = 2$ ,  $p < .001$ ). Die unstandardisierten Ladungen waren  $\lambda_3 = 0$ ,  $\lambda_6 = 1$  und  $\lambda_9 = 1.15$  für Mathematik sowie  $\lambda_3 = 0$ ,  $\lambda_6 = 1$  und  $\lambda_9 = 1.37$  für Deutsch. Der Lernzuwachs wird also insgesamt flacher, insbesondere im Fach Mathematik (siehe dazu Moser et al., 2017). Sowohl das Ausgangsniveau der Schulleistungen als auch der Zuwachs hatten signifikante Varianz ( $p < .001$ ), was eine wichtige Voraussetzung für die folgenden Analysen darstellt. Die Korrelation zwischen den beiden Wachstumskomponenten belief sich auf  $r = -.17$  ( $p = .004$ ), was darauf hinweist, dass zunächst leistungsstärkere Schüler etwas langsamere Lernzuwächse

aufweisen und es somit zu einer geringfügigen Homogenisierung der Schulleistungen über die obligatorische Schulzeit hinweg kommt. Die aufgeklärte Varianz in dem beobachteten Variablen ist mit  $.71 < R^2 < .85$  sehr hoch, das Modell kann also sehr gut die interindividuellen Unterschiede in den Verläufen der Schulleistung abbilden.

Ausgehend von dem Messmodell wurden anschließend Mischverteilungsmodelle mit einer zunehmenden Anzahl von Typen berechnet (siehe Wang & Bodner, 2007). Als Hilfsvariablen (AUXILIARY-Befehl; siehe Asparouhov & Muthén, 2014) für die Schätzung der Typenzugehörigkeit wurden die drei Skalen des schulbezogenen Vorwissens (siehe unten) verwendet. Die numerische Integration basierte auf 10.000 Startwerten und eine Lösung wurde nur dann akzeptiert, wenn sie über mehrere Startwerte hinweg konvergierte. Damit die Modelle konvergierten, musste angenommen werden, dass die Varianzen der Wachstumskomponenten über alle Typen hinweg gleich waren. Die Bestimmung der optimalen Anzahl der Typen erfolgte unter Berücksichtigung von vier Indikatoren, die gesamthaft betrachtet wurden: die Log-Likelihood-Differenz zu Modell mit  $k-1$  Typen ( $2\Delta LL$ ; siehe Lo, Mendell & Rubin, 2001), das Informationskriterium nach Akaike (AIC; siehe Akaike, 1987; Burnham, Anderson & Huyvaert, 2011), das Informationskriterium nach Bayes ohne und mit Korrektur für die Stichprobengröße (BIC/SSA-BIC; siehe Schwartz, 1978; Sclove, 1987) sowie das Entropiemaß (siehe Ramaswamy, DeSarbo, Reibstein & Robinson, 1993). Mit der Log-Likelihood-Differenz wird geprüft, ob ein Modell mit  $k$  Typen zu einer signifikant besseren Modellpassung führt als ein Modell mit  $k-1$  Typen. Die Informationskriterien zeigen die absolute Modellpassung an und berücksichtigen dabei zum Teil die Modellkomplexität, wobei kleinere Werte auf eine bessere Passung hindeuten. Das Entropiemaß ist ein standardisiertes Maß der Klassifikationsgüte und kann Werte zwischen 0 und 1 annehmen. Höhere Werte deuten auf eine bessere Klassifikation der Individuen zu den Typen auf Grundlage der individuellen *a posteriori*-Wahrscheinlichkeiten.

#### **4.3 Operationalisierung von Variablen auf der Schülerebene**

Auf der Schülerebene wurden das Geschlecht, das Alter zum Zeitpunkt der Einschulung, die psychometrische Intelligenz, das schulbezogene Vorwissen sowie das Arbeits- und Sozialverhalten erfasst und als Prädiktoren für die Typenzugehörigkeit verwendet. Die deskriptiven Statistiken zu diesen Variablen finden sich im oberen Teil der Tabelle 4.

**Geschlecht.** Die Angaben zum Geschlecht der Schülerinnen und Schüler entstammen der Bildungsstatistik BISTA des Kantons Zürich.

**Alter zum Zeitpunkt der Einschulung.** Die Angaben zum Alter der Schülerinnen und Schüler entstammen der Bildungsstatistik BISTA des Kantons Zürich.

**Zeitpunkt der Einschulung.** Zusätzlich zum Alter wurde aus der Bildungsstatistik des Kantons Zürich bestimmt, ob die Schülerin oder der Schüler zum normativen Zeitpunkt (87.6%) eingeschult wurde, ob die Einschulung um ein Schuljahr zurückgestellt wurde (9.4%) oder ob eine vorzeitige Einschulung stattfand (3.0%). Eine Rückstellung kann von der Schulbehörde angeordnet werden und eine vorzeitige Einschulung von der Schulbehörde bewilligt werden, sofern die Umstände diese Maßnahme als geboten erscheinen lassen. Für die Auswertungen wurde die Variable dummykodiert, wobei die normative Einschulung als Referenzkategorie verwendet wurde.

**Psychometrische Intelligenz.** Die allgemeinen kognitiven Fähigkeiten wurden zum ersten Erhebungszeitpunkt beim Schuleintritt (T1) mit dem Culture Fair Test (CFT 1; Weiß & Osterland, 1997) erfasst. Der CFT 1 erfasst nonverbale, kulturell unabhängige Intelligenz und besteht aus den Untertests „Substitutionen“, „Labyrinthe“, „Klassifikationen“, „Ähnlichkeiten“ und „Matrizen“. Bezüglich der konvergenten Validität zeigen sich mittlere Korrelationen zwischen der CFT 1-Gesamtleistung und dem HAWIK (Handlungsteil  $r = .66$  und Verbalteil  $r = .48$ ; Weiß & Osterland, 1997). Der Mittelwert in der Ausgangsstichprobe betrug  $M = 100.36$  ( $SD = 14.90$ ).

**Schulbezogenes Vorwissen.** Das schulbezogene Vorwissen in den drei Aspekten

„Lesekompetenz“, „Wortschatz“ und „Rechenkompetenz“ wurde mit Hilfe des Tests „Sprache und Mathematik beim Schuleintritt“ (SMS; Moser, Berweger & Tresch, 2003) erfasst. Der Test wurde in Einzelausschüttungen durchgeführt und dauerte zwischen 10 und 20 Minuten. Die Lesekompetenz wurde mit den Subtests „Buchstaben benennen“, „Silben lesen“, „Wörter lesen und verstehen“ sowie „Sätze lesen und verstehen“ gemessen. Die Subtests „Aktiver Wortschatz: Nomen“ und „Aktiver Wortschatz: Verben“ dienten zur Messung des Wortschatzes. Beim Rechnen wurden der Kardinalzahlaspekt durch den Subtest „Zählen und Zahl zuordnen“, der Ordinalzahlaspekt durch den Subtest „Reihen bilden“ und der Rechenaspekt durch den Subtest „Addition und Subtraktion“ abgedeckt. Beim Rechnen wurde außerdem der Subtest „Zahlsymbole benennen“ durchgeführt. Weitere Details zum Test sowie Itembeispiele finden sich bei Moser, Berweger und Stamm (2005a, 2005b, 2005c). Die Skalierung erfolgte analog zu der Skalierung der Leistungstests, mit der Ausnahme, dass die Punktwerte in den drei Skalen zum schulbezogenen Vorwissen nicht mit den Punktwerten in den Leistungstests gelinkt wurden. Stattdessen wurden die Mittelwerte der Skalen willkürlich auf  $M = 500$  ( $SD = 100$ ) Punkte festgelegt.

**Arbeits- und Sozialverhalten.** Das Arbeits- und Sozialverhalten der Schülerinnen und Schüler wurden durch eine Einschätzung der Lehrpersonen mit Hilfe der *School Social Behavior Scales* von Merrell (1999) bestimmt und auf vier Subskalen abgebildet. Zur Erfassung des *regelkonformen Verhaltens* wurden Verhaltensweisen beurteilt, die sich auf die Einhaltung der durch die Lehrperson oder der gemeinsam mit der Klasse festgelegten Regeln beziehen. Beispielitems sind „Die Schülerin, der Schüler folgt den Regeln im Schulzimmer“ oder „Die Schülerin, der Schüler erfüllt zugewiesene Anordnungen der Lehrperson und führt diese aus“. Die Skala zum *selbständigen Verhalten* ist konzeptuell mit der Skala des regelkonformen Verhaltens verwandt, bezieht sich jedoch stärker auf das Ausführen und Bearbeiten von Arbeitsaufträgen im Unterrichtskontext. Beispielitems sind „Die Schülerin, der Schüler erfüllt zugewiesene Aufgaben selbstständig“ oder „Die Schülerin, der Schüler

macht individuelle Stillarbeiten fertig, ohne dazu aufgefordert zu werden“. Als drittes Merkmal des Arbeits- und Sozialverhaltens wird *kooperatives Verhalten* erfasst. Es zielt auf das Verhalten während Partner- und Gruppenarbeiten und wird mit Items wie „Die Schülerin, der Schüler bietet Mitschülerinnen und Mitschülern wenn nötig Hilfe an“ oder „Die Schülerin, der Schüler ermuntert andere Schülerinnen und Schüler zur Beteiligung an Aktivitäten“. Die letzte Skala zur *sozialen Integration* erfasst das Arbeits- und Sozialverhalten indirekt, indem sie misst, wie Mitschülerinnen und Mitschüler auf das Kind reagieren. Items zur Erfassung der sozialen Integration sind beispielsweise „Die Schülerin, der Schüler wird von den Mitschülerinnen und Mitschülern in gemeinsame Aktivitäten einbezogen“ und „Die Schülerin, der Schüler wird von anderen Mitschülerinnen und Mitschülern ausgelacht“. Weitere Informationen zu dem eingesetzten Instrument und dem Zusammenhang von Arbeits- und Sozialverhalten mit soziodemographischen Merkmalen finden sich bei Tresch (2005). Die vier postulierten Skalen konnten mit Hilfe einer Hauptkomponentenanalyse in den Daten repliziert werden und die internen Konsistenzen nach Cronbach sind mit  $.78 < \alpha < .89$  als zufriedenstellend bis gut zu bezeichnen.

---

#### TABELLE 4 ETWA HIER

---

#### 4.4 Operationalisierung der Variablen auf der Ebene des Herkunftskontexts

Auf der Ebene des Herkunftskontexts wurden die Herkunftssprache, der Bildungshintergrund der Eltern sowie die vorschulischen Freizeitaktivitäten und der vorschulische Medienkonsum erfasst. Eine deskriptive Zusammenfassung dieser Variablen findet sich im unteren Teil von Tabelle 4.

**Herkunftssprache.** Diese unterscheidet zwischen Kindern mit Deutsch als Erstsprache und solchen mit Deutsch als Zweitsprache, wobei kein Unterschied zwischen den deutschen Standardsprachen (Schweizerhochdeutsch oder Hochdeutsch) oder Dialekten

gemacht wurde. Die Erstsprache wurde als Sprache erfasst, welche die Eltern am häufigsten mit ihrem Kind sprechen. Bei Kindern mit Deutsch als Erstsprache entspricht die Erstsprache auch der Unterrichtssprache. Die Angaben stammen größtenteils aus dem Elternfragebogen zum Testzeitpunkt T2. Damit auch Eltern mit geringen Sprachkenntnissen erreicht werden konnten, wurde der Elternfragebogen in sieben Sprachen übersetzt. Die Rücklaufquote betrug 73.0%, fehlende Daten zur Erstsprache wurden durch Daten der Elternbefragung beim Schuleintritt (T1) oder durch Daten der Schülerinnen- und Schülerbefragung am Ende der 6. (T3) bzw. am Ende der 9. Klasse (T3) ergänzt. Zwei Drittel (72.4%) der Kinder sprachen demnach Deutsch als Erstsprache und ein Drittel (27.6%) als Zweitsprache.

**Bildungshintergrund.** Der Bildungshintergrund der Schülerinnen und Schüler wurde über den höchsten Schulabschluss der Eltern operationalisiert. Diese Angaben stammten zum größten Teil aus dem Elternfragebogen zu T1 oder, wenn sie dort fehlten, aus den Auskünften der Schülerinnen und Schüler zu späteren Messzeitpunkten. Die kategorialen Angaben (z. B. „Gymnasium“) wurden in die Dauer der schulischen und beruflichen Bildung in Jahren umgerechnet. Wenn von beiden Eltern eine Angabe vorlag, wurde die jeweils längere ausgewählt. Im Durchschnitt berichteten die Eltern bzw. ihre Kinder über eine Bildungsdauer von  $M = 12.58$  ( $SD = 3.52$ ) Jahren.

**Freizeitaktivitäten.** Im Elternfragebogen wurden die Eltern gefragt, wie oft ihre Kinder an einer Auswahl von strukturierten Freizeitaktivitäten (z. B. Musikunterricht, Sportverein, Fremdsprachenunterricht oder Computerkurs) und gemeinsamen Familienaktivitäten (z. B. gemeinsam malen, gemeinsam lesen, gemeinsam Spiele machen oder gemeinsame Aktivitäten im Freien) teilnehmen würden. Insgesamt lagen 18 Items vor, die auf einer vierstufigen Skala von „nie“ bis „mehrmals pro Woche“ (für strukturierte Freizeitaktivitäten) bzw. von „nie oder selten“ bis „täglich“ (für gemeinsame Familienaktivitäten) beantwortet werden konnten. Eine explorative Faktorenanalyse dieser 18 Items ergab eine zweifaktorielle Lösung, welche varimaxrotiert weitgehend die Struktur der

außerfamiliären strukturierten und familiären unstrukturierten Aktivitäten widerspiegelte. Die drei höchsten Ladungen für die außerfamiliären Freizeitaktivitäten gab es für „Computerkurs“ (.72), „Begabungsförderung“ (.72) und „Deutschkurs“ (.54). Die drei höchsten Ladungen für familiäre Freizeitaktivitäten gab es für „gemeinsam lesen/Geschichten hören“ (.71), „Geschichten vorlesen“ (.64) und „gemeinsam Spiele machen“ (.57). Für die weiteren Auswertungen wurden die zwei Faktorwerte mit der Methode von Bartlett bestimmt.

**Mediennutzung.** Zwei Items im Elternfragebogen bezogen sich auf die durchschnittliche Dauer des täglichen Fernsehkonsums (auf einer fünfstufigen Skala von „nie“ bis „mehr als 2 Stunden pro Tag“) und die Häufigkeit der Computernutzung (auf einer vierstufigen Skala von „nie“ bis „täglich“). Die beiden Items waren weitestgehend unkorreliert ( $r = .04$ ;  $p = .079$ ) und wurden im Folgenden als Einzelitemindikatoren verwendet.

### 4.3 Stichprobenausfall

Um abzuschätzen, ob sich die Zusammensetzung der Stichprobe durch Wegzug aus dem Kanton Zürich oder durch selektive Nichtteilnahme verändert hatte, wurden die nicht veränderlichen Merkmale, also Geschlecht, Alter zum Zeitpunkt der Einschulung, Fremdsprachigkeit und Bildungshintergrund der Eltern über die vier Messzeitpunkte hinweg betrachtet. Die Ergebnisse dieser Betrachtung sind in Tabelle 5 zusammengestellt und legen nahe, dass die Nichtteilnahme die Zusammensetzung der Stichprobe nur unwesentlich verändert hat. Dieser Befund spricht für eine hohe Generalisierbarkeit der folgenden Befunde.

=====

TABELLE 5 ETWA HIER

=====

## 5 Ergebnisse

### 5.1 Bestimmung der optimalen Zahl der Verlaufstypen

Die unterschiedlichen Indikatoren zur Bestimmung der optimalen Anzahl der



Verlaufstypen ergeben kein einheitliches Bild (siehe Tabelle 6). Der Log-Likelihood-Test nach Lo et al. (2001) zeigt, dass zwei Typen zu einer besseren Modellpassung führen als ein Typ, und dann wieder, dass fünf Typen besser sind als vier. Der AIC sinkt kontinuierlich mit einer steigenden Zahl von Typen an und erreicht bis zur Sechs-Typen-Lösung immer noch kein Minimum. Der BIC hat sein Minimum bei vier Klassen, während der stichprobengrößenkorrigierte BIC sein Minimum bei fünf Klassen hat. Die Entropie sinkt zunächst mit steigender Anzahl der Verlaufstypen, steigt dann wieder bei der Fünf-Typen-Lösung an, um dann wieder leicht abzusinken. Modelle mit mehr als sechs Typen liefern zum Teil wieder bessere Anpassungswerte, sind aber inhaltlich wenig nützlich, da nur noch sehr kleine „Spezialtypen“ mit wenigen Schülerinnen und Schülern identifiziert werden. Gesamthaft betrachtet kommen also Modelle mit zwei, vier und fünf Typen in Frage, wobei die meisten Indikatoren für das Fünf-Typen-Modell sprechen, das wir aus diesem Grund für die folgenden Auswertungen beibehalten haben.

=====

TABELLE 6 ETWA HIER

=====

## 5.2 Beschreibung der Verlaufstypen

Die fünf unterschiedlichen Verlaufstypen sind in der Stichprobe unterschiedlich stark vertreten und zeichnen sich, wie erwartet, durch unterschiedliche Verläufe aus, die im Folgenden skizziert werden sollen. Eine Zusammenfassung findet sich auch in Tabelle 7. Etwa ein Fünftel der Schülerinnen und Schüler (20.1%) konnte dem *durchschnittlichen Verlaufstyp* zugeordnet werden. Sein Ausgangsniveau liegt bei 472 Punkten und der durchschnittlicher Lernzuwachs bei 124 Punkten pro Schuljahr in der Primarschule. Ein etwas tieferes Ausgangsniveau und einen etwas langsameren Lernzuwachs kann dem *lernschwachen Verlaufstyp* zugeordnet werden. Er umfasst ein gutes Viertel der Stichprobe (26.3%), hat ein Ausgangsniveau von 451 Punkten und einen Lernzuwachs von nur 116 Punkten. Am oberen

Ende der Leistungsverteilung findet sich der *Verlaufstyp mit Vorsprung*, der zwar mit einem jährlichen Leistungszuwachs von 122 Punkten eher durchschnittlich ist, dafür aber ein sehr hohes Anfangsniveau von 617 Punkten aufweist. Dieses hohe Anfangsniveau liegt fast 150 Punkte höher als das Anfangsniveau beim durchschnittlichen Verlaufstyp und übersetzt sich in mehr als ein Schuljahr Vorteil gegenüber dieser Gruppe. Der Verlaufstyp mit Vorsprung umfasst 14.6% der Schülerinnen und Schüler.

Theoretisch besonders interessant sind die verbleibenden zwei Gruppen. Mit 30.1% stellt der *ertragsreiche Verlaufstyp* zahlenmäßig den größten Anteil an allen Schülerinnen und Schülern. Er zeichnet sich einerseits durch ein leicht überdurchschnittliches Anfangsniveau von 512 Punkten aus, doch bemerkenswerter ist der jährliche Lernzuwachs von 142 Punkten. Dieser Typ profitiert beim jährlichen Lernzuwachs etwa 15% mehr als der durchschnittliche Typ und 22% mehr als der lernschwache Typ, was sich über die Schulzeit so weit aufaddiert, dass er das gleiche Niveau erreicht wie der Typ mit Vorsprung. Die restlichen etwa 9% können als der *nachzügeln*de Verlaufstyp bezeichnet werden und sind aus der Perspektive der Leistungsentwicklung als besonders problematisch anzusehen. Dieser Typ weist mit 436 Punkten nicht nur das tiefste Ausgangsniveau auf, sondern ist mit einem jährlichen Lernzuwachs von nur 106 Punkten auch der langsamste und holt das tiefe Ausgangsniveau auch nicht mehr auf.

---

#### TABELLE 7 ETWA HIER

---

Ein weiteres Unterscheidungsmerkmal zwischen den Verlaufstypen sind *relative* Leistungsunterschiede zwischen den Fächern Deutsch und Mathematik, die sich statistisch als Mittelwertunterschied auf der ersten Ebene des Wachstumskurvenmodells widerspiegeln und in den beiden rechten Spalten von Tabelle 7 dargestellt sind. Dabei lassen sich verschiedene Beobachtungen machen. So fällt auf, dass vor allen Dingen der ertragsreiche, aber auch der

lernschwache Typ mit relativ starken Leistungen in Deutsch (bzw. relativ schwachen Leistungen in Mathematik) startet, also so etwas wie eine „sprachliche Begabung“ hat. Dem gegenüber sind der durchschnittliche Verlaufstyp und der Typ mit Vorsprung relativ stark in Mathematik (bzw. relativ schwach in Deutsch). Schaut man sich die relativen Lernzuwächse an, so fällt vor allem auf, dass der ertragsreiche Typ, aber auch der lernschwache Typ relativ gesehen geringere Zuwächse in den Deutschleistungen (bzw. relativ gesehen höhere Zuwächse in Mathematik) zeigt. Der durchschnittliche Typ sowie der Typ mit Vorsprung dagegen wachsen relativ stark im Fach Deutsch (bzw. relativ schwach im Fach Mathematik). Stellenweise sind die Effektstärken in den relativen Leistungsunterschieden substantiell.

### **5.3 Korrelate der Zugehörigkeit zu einem Verlaufstyp**

In Tabelle 8 finden sich die Korrelationen zwischen der Wahrscheinlichkeit der Zugehörigkeit zu jedem der fünf Verlaufstypen auf der einen Seite mit Variablen auf der Schülerebene auf der anderen. Zunächst einmal fällt auf, dass die Koeffizienten vor allem für die beiden „Extremtypen“ der nachzügeln den bzw. der bevorteilten Schülerinnen und Schüler besonders hoch sind, während sie für die drei „Mitteltypen“ weniger stark ausgeprägt sind. Die psychometrische Intelligenz sowie die drei Aspekte des vorschulischen Wissens scheinen diejenigen Variablen zu sein, welche zwischen den fünf Verlaufstypen am besten differenzieren. Unter den vier Variablen zur Erfassung des Arbeits- und Sozialverhaltens scheint die Selbständigkeit eine hervorragende Rolle zu spielen. Während das Alter und der Zeitpunkt der Einschulung weniger relevant zu sein scheinen, finden sich für das Geschlecht kleinere Effekte.

=====

TABELLE 8 ETWA HIER

=====

In Tabelle 9 finden sich die gleichen Zusammenhänge für Variablen auf der Ebene des Herkunftskontextes. Während der Bildungshintergrund der Eltern und die Herkunftssprache

für die Vorhersage aller fünf Verlaufstypen relevant sind, zeigen sich beim Freizeitverhalten weniger systematische Effekte. Lediglich die TV-Nutzung scheint einen Effekt auf die Typenzugehörigkeit zu haben und differenziert zwischen dem nachzügeln, lernschwachen und durchschnittlichen Verlaufstyp einerseits und dem ertragsreichen sowie dem Typ mit Vorsprung andererseits.

---

---

TABELLE 9 ETWA HIER

---

---

#### 5.4 Multivariate Vorhersage der Zugehörigkeit zu einem Verlaufstyp

Um für wechselseitige Einflüsse die betrachteten Variablen zu kontrollieren und eine möglichst sparsame Vorhersage der Zugehörigkeit zu einem Verlaufstyp zu erhalten, wurden die fünf Wahrscheinlichkeiten der Typenzugehörigkeit mit einer STEPWISE-Regression vorhergesagt, wobei als Einschlusskriterium eine Irrtumswahrscheinlichkeit von  $p < .001$  gewählt wurde. Die so identifizierten Regressionskoeffizienten werden in der Reihenfolge der Aufnahme in die Regressionsgleichung berichtet.

Die Zugehörigkeit zum nachzügeln Typ konnte durch neun Variablen vorhergesagt werden, nämlich psychometrische Intelligenz ( $\beta = -.15$ ), Wortschatz ( $\beta = -.14$ ), Alter zum Zeitpunkt der Einschulung ( $\beta = .11$ ), selbständiges Arbeitsverhalten ( $\beta = -.13$ ), Teilnahme an strukturierten Freizeitaktivitäten ( $\beta = .10$ ), Bildungsabschluss der Eltern ( $\beta = -.09$ ), soziale Integration ( $\beta = .10$ ), Geschlecht ( $\beta = -.12$ , so dass Mädchen in diesem Verlaufstyp seltener vertreten waren) sowie schließlich Rechenleistungen ( $\beta = -.18$ ). Schwache Rechenleistungen, eine niedrige psychometrische Intelligenz und ein geringer Wortschatz sind die drei Variablen mit den stärksten unigen Effekten. Das Modell erklärt (korrigiert)  $R^2 = .24$  der Varianz in dieser Typenzugehörigkeitsvariable.

Die Zugehörigkeit zum lernschwachen Typ konnte nur durch sechs Variablen vorhergesagt werden, wobei die aufgeklärte Varianz mit  $R^2 = .14$  eher gering ausfällt. In der

Reihenfolge der Aufnahme in die Regressionsgleichung sagen die psychometrische Intelligenz ( $\beta = -.14$ ), die Rechenleistungen ( $\beta = -.11$ ), die Bildung der Eltern ( $\beta = -.11$ ), das Geschlecht ( $\beta = .13$ , es sind also häufiger Mädchen vertreten), die Selbstständigkeit ( $\beta = -.10$ ) sowie eine häufigere TV-Nutzung ( $\beta = .09$ ) die Zugehörigkeit zu diesem Verlaufstyp vorher. Zusammengefasst handelt es sich hier also um weniger intelligente Mädchen mit weniger privilegiertem sozialem Hintergrund.

Auch die Zugehörigkeit zum durchschnittlichen Verlaufstyp konnte nur durch vier Variablen signifikant vorhergesagt werden, wobei hier nur  $R^2 = .13$  der Varianz aufgeklärt werden konnte. Im durchschnittlichen Verlaufstyp fanden sich eher Kinder mit schwächeren Leseleistungen ( $\beta = -.23$ ), häufiger Jungen ( $\beta = -.18$ ), Kinder mit schlechterem Wortschatz ( $\beta = -.09$ ) und Kinder, die selbst später eingeschult wurden ( $\beta = .10$ ). Wir haben es also eher mit Jungen mit geringeren vorschulischen Kenntnissen zu tun, die aber in ihrer Intelligenz oder ihrem sozialen Hintergrund nicht über- oder unterdurchschnittlich sind.

Die Zugehörigkeit zum ertragsreichen Typ konnte durch sechs Variablen erklärt werden, wobei die aufgeklärte Varianz mit  $R^2 = .21$  etwas höher war. In diesem Verlaufstyp finden sich Lernende mit überdurchschnittlichen Leseleistungen ( $\beta = .18$ ), eher Mädchen ( $\beta = .19$ ) mit überdurchschnittlich gebildeten Eltern ( $\beta = .12$ ), Kinder, die jünger bei der Einschulung waren ( $\beta = -.16$ ), und solche, die über einen größeren Wortschatz ( $\beta = .12$ ) und über ein selbständigeres Arbeitsverhalten verfügten ( $\beta = .12$ ). Insgesamt drängt sich also das Bild von lesefreudigen Mädchen bildungsinteressierter Eltern auf.

Die Zugehörigkeit zum Typ mit Vorsprung schließlich konnte mit fünf Variablen bei einer aufgeklärten Varianz von  $R^2 = .30$  recht gut vorhergesagt werden. Die größte Relevanz hatten hier die Rechenleistungen ( $\beta = .26$ ) sowie die psychometrische Intelligenz ( $\beta = .25$ ). Außerdem spielen für die Zugehörigkeit zu diesem Verlaufstyp der Bildungshintergrund der Eltern ( $\beta = .11$ ), das Geschlecht der Kinder ( $\beta = -.11$ , so dass etwas weniger Mädchen in diesem Verlaufstyp zu finden sind) sowie das selbständige Arbeitsverhalten ( $\beta = .09$ ) eine

Rolle. Wir haben es also hier mit rechenstarken und überdurchschnittlich intelligenten Kindern zu tun, die etwas häufiger Jungen und nicht Mädchen sind.

## **6 Diskussion**

Erstmalig für den deutschsprachigen Bildungsraum liegen Befunde vor, die aufzeigen, dass sich die langfristigen individuellen Lernverläufe von Schülerinnen und Schülern in unterschiedliche Verlaufstypen gruppieren lassen. In der Diskussion dieser Befunde möchten wir zunächst die wichtigsten Ergebnisse zusammenfassen, bevor wir auf die Stärken und Schwächen der Studie eingehen und die theoretischen und praktischen Implikationen der Befunde aufzeigen. Abschließend werden wir offen gebliebene Fragen ansprechen und einen Ausblick auf die zukünftige Forschung wagen.

### **6.1 Zusammenfassung der Ergebnisse**

Die Analyse der vorliegenden längsschnittlichen Daten einer großen Zufallsstichprobe von Schülerinnen und Schülern zeigt auf, dass sich die Verläufe von Schulleistungen in fünf qualitativ unterschiedliche Verlaufstypen kategorisieren lassen. Für die Modellierung des Verlaufs der Schulleistungen wurde ein Wachstumsmodell zweiter Ordnung gewählt, wobei auf der Ebene der beiden Schulfächer eine Ladung frei geschätzt wurde. Dieses Modell wurde gewählt, um erstens möglichen nichtlinearen Leistungsverläufen (siehe Bloom et al., 2008) Rechnung zu tragen und zweitens die Entwicklung der Schulleistung insgesamt und nicht bezogen auf fächerspezifische Aspekte, zu denen wir keine Hypothesen hatten, abzubilden.

Für die Bestimmung der Anzahl der unterschiedlichen Verlaufstypen wurde eine Reihe von statistischen Kennzahlen verwendet, die mehrheitlich auf eine Lösung mit fünf distinkten Typen hinwiesen. Diese Zahl ist einerseits ausreichend groß, um die erhebliche Heterogenität in den Leistungsverläufen widerzuspiegeln und um gezielte bildungspolitische oder pädagogische Interventionen zu begründen. Sie ist andererseits aber auch ausreichend klein, um eine genügend große Besetzung der Gruppen zu gewährleisten und damit praktisch bedeutsam zu sein. Lösungen mit einer geringeren Anzahl von Gruppen lieferten triviale

Lösungen, während Lösungen mit einer höheren Anzahl von Gruppen zusätzlich nur noch sehr kleine und sehr spezielle Verlaufstypen identifizierten. Die Lösung mit fünf Gruppen war also nicht nur unter statistischen, sondern auch unter inhaltlichen und Anwendungsgesichtspunkten ideal.

Absteigend sortiert nach der Besetzungshäufigkeit identifizierten wir den ertragsreichen Verlaufstyp (30.1%), den lernschwachen Verlaufstyp (26.3%), den durchschnittlichen Verlaufstyp (20.1%), den Verlaufstyp mit Vorsprung (14.6%) sowie den nachzügeln den Verlaufstyp (8.9%). Auffällig ist dabei der große Anteil des ertragsreichen Verlaufstyps, der nicht etwa einen außergewöhnlichen und seltenen, sondern tatsächlich den häufigsten Verlaufstyp darstellt. Gegenüber dem durchschnittlichen Verlaufstyp zeichnet sich der ertragsreiche Verlaufstyp durch einem im Durchschnitt 15% schnelleren Lernzuwachs aus, was sich über die gesamte Schulzeit zu einem beträchtlichen Vorsprung addiert. Pädagogisch interessant ist hier die Identifikation von Merkmalen, die zu so einem starken Lernzuwachs führt.

Dem ertragsreichen Verlaufstyp ist von der zahlenmäßigen Besetzung her und aus inhaltlichen Überlegungen der lernschwache Verlaufstyp gegenübergestellt. Gegenüber dem durchschnittlichen Verlaufstyp weist er einen um 6% langsameren Lernzuwachs auf, der sich im Verlauf der gesamten Schulzeit zu einem beträchtlichen Nachteil kumuliert. Hier stellt sich die Frage, ob sich dieser Typ dadurch auszeichnet, dass ihm die Merkmale des ertragsreichen Typen fehlen, oder ob unabhängig davon Risikofaktoren vorhanden sind, die zu einem geringen Lernzuwachs führen. Zusammengenommen machen diese beiden „Haupttypen“ mehr als die Hälfte und einschließlich des durchschnittlichen Typs etwa drei Viertel der Stichprobe aus.

Daneben konnten wir in den Daten zwei seltenere, aber inhaltlich umso interessantere Verlaufstypen identifizieren. Den Typ mit Vorsprung mit einem erheblichen Ausgangsniveau, aber nur durchschnittlichem Leistungszuwachs, sowie den nachzügeln den Verlaufstyp mit

unterdurchschnittlichem Ausgangsniveau und unterdurchschnittlichem Lernzuwachs. Bei dem Typen mit Vorsprung stellt sich vor allem die Frage, warum diese Schülerinnen und Schüler trotz ihres hohen Ausgangsniveaus nicht deutlich mehr vom Unterricht profitieren, während beim nachzügeln Typ die Identifikation der stärksten und damit vermutlich ausschlaggebenden Korrelate der Zugehörigkeit zu diesem Verlaufstyp im Mittelpunkt des Interesses steht.

Bevor wir versuchen, die Antworten auf die aufgeworfenen Fragen zu finden, sei hervorgehoben, dass diese Fragen ohne das personenzentrierte Vorgehen bei der Identifikation der fünf Verlaufstypen gar nicht gestellt werden könnten. Ohne den Einsatz von Mischverteilungsmodellen wäre weder die Identifikation der unterschiedlichen Typen noch die Bestimmung nichtlinearer Prädiktoren der Typenzugehörigkeit möglich. Im besten Fall hätte man die Varianz im Ausgangsniveau und die Varianz im Verlauf identifizieren und vorhersagen können, ohne dass distinkte Kombinationen in den Verlaufsparemtern (wie etwa die eines starken Anstiegs in Kombination mit einem durchschnittlichen Ausgangsniveau wie bei dem ertragsreichen Typ) hätten gemeinsam betrachtet werden können. Mit dem personenzentrierten Vorgehen wird man also nicht nur der Heterogenität der Leistungsverläufe besser gerecht, sondern muss auch nicht mehr von der impliziten (und wahrscheinlich unrealistischen) Annahme ausgehen, die untersuchten Zusammenhänge und Prozesse seien für alle untersuchten Personen gleich (vgl. Magnusson, 2003). Und tatsächlich deuten die Befunde darauf hin, dass die fünf Verlaufstypen nicht entlang einer Dimension angeordnet sind, was die jeweiligen Prädiktorvariablen angeht. Vielmehr zeigen sich, insbesondere in der multivariaten Betrachtung, unterschiedliche Konstellationen von Prädiktoren, die alles in allem ein sehr schlüssiges Bild ergeben.

Nach dieser Feststellung wollen wir versuchen, uns den eben aufgeworfenen Fragen zu nähern, um damit die wichtigsten Befunde zusammenzufassen. Wenn man sich zunächst die Merkmale des ertragsreichen Verlaufstyps anschaut, so fällt auf, dass sprachliche



Fähigkeiten (operationalisiert über die Leseleistungen und den Wortschatz bei der Einschulung) hier im Vordergrund stehen. Auch sind die mathematischen Fähigkeiten in diesem Verlaufstyp relativ gesehen zu den sprachlichen Fähigkeiten am geringsten von allen fünf Verlaufstypen ausgeprägt, was die Bedeutsamkeit der sprachlichen Fähigkeiten zusätzlich hervorhebt. Wenn man davon ausgeht, dass jeder Unterricht Sprache voraussetzt und auf Sprache aufbaut, verwundert es nicht, dass Kinder mit überdurchschnittlichen sprachlichen Fähigkeiten auch überdurchschnittlich vom Unterricht profitieren. Das tun sie auch (und vielleicht insbesondere) in den nichtsprachlichen Fächern, was sich in den relativ hohen Mathematikleistungszuwächsen in dieser Gruppe zeigt. Der zweite Faktor neben der Sprache scheint eine Passung der persönlichen Merkmale zu den Anforderungen von Schule und Unterricht zu sein (vgl. etwa Helsper, Kramer & Thiersch, 2014). Mädchen, die ja häufiger schulisch angemessenes und angepasstes Verhalten zeigen, selbständigere Kinder und solche, die über das bildungsnähere Elternhaus wahrscheinlich mehr Wertschätzung für die Bildungsthematik erfahren haben, sind in diesem Verlaufstyp überproportional vertreten. Aus der Perspektive eines Angebot-Nutzungs-Modells (siehe Helmke, 2015) sind demnach – zumindest in dieser Stichprobe – zwei Variablengruppen identifiziert, die eine besonders effiziente Nutzung vorhersagen: Sprachkompetenz und Anpassungsfähigkeit. Dagegen ist die psychometrische Intelligenz dieser Kinder, entgegen aller Plausibilität, nicht auffällig hoch ausgeprägt.

Bemerkenswerterweise lässt sich der lernschwache Typ, sozusagen als Pendant zu dem ertragsreichen Typ, eben *nicht* als dessen Gegenpol vorhersagen, denn auch hier sind beispielsweise Mädchen überrepräsentiert und die sprachlichen Fähigkeiten spielen zur Vorhersage der Typenzugehörigkeit keine große Rolle. Was bei diesen Kindern dagegen geringer ausgeprägt ist, sind die psychometrische Intelligenz und, möglicherweise damit zusammenhängend, die Rechenleistungen. Die insgesamt niedrige Kraft zur Vorhersage dieses Verlaufstyps durch die untersuchten Variablen spricht außerdem dafür, dass hier

wesentlich individuellere Faktoren (wie etwa Schwierigkeiten in der Familie oder eine Verhaltensauffälligkeit) eine größere Rolle spielen als die verwendeten breiten Marker von Person und Herkunft.

Der Zugehörigkeit zum Verlaufstyp mit Vorsprung ließ sich dagegen relativ gut vorhersagen, und hier waren es vor allen Dingen die hohe psychometrische Intelligenz und, möglicherweise damit zusammenhängend, die hohen Rechenleistungen bereits kurz nach der Einschulung. Außerdem ist der Bildungshintergrund eher privilegiert und das Arbeitsverhalten selbständiger als im Durchschnitt. Warum diese Kinder bei diesen Voraussetzungen nicht noch mehr vom Unterricht profitieren, und zwar auch nicht vom Mathematikunterricht, wie die Analyse der Residuen nahelegt, lässt sich an dieser Stelle nur spekulieren. Vielleicht sind es am Ende doch die lediglich durchschnittlichen sprachlichen Fähigkeiten, die dafür verantwortlich sind, dass diese Kinder in Durchschnitt genauso viel vom Unterricht mitnehmen wie andere Kinder im Durchschnitt auch. Ohne Zweifel starten die Kinder dieses Verlaufstyps sehr stark und schaffen es auch, diesen Startvorteil über die gesamte obligatorische Schulzeit aufrechtzuerhalten, so dass sie dann am Ende auch mit den höchsten Schulleistungen die obligatorische Schule verlassen. Aber gerade die Kinder vom ertragsreichen Typ zeigen, dass deutlich höhere Lernzuwächse ja auch möglich wären. Alternativ zu der Erklärung über die durchschnittlichen sprachlichen Fähigkeiten kann man darüber spekulieren, ob diesen Kindern in dem gegebenen Schulsystem nicht genug Fördermöglichkeiten offenstehen, um ihren Begabungen (und gegebenenfalls schon ihrer Hochbegabung) adäquat gerecht zu werden.

Zum Schluss bleibt die Betrachtung des nachzügeln den Verlaufstyps, und hier stellt sich die Frage, ob es vielleicht einzelne vorhersagestarke Faktoren gibt, welche die Zugehörigkeit zu diesem Verlaufstyp vorhersagen könnten. Diese Frage muss an dieser Stelle ganz klar verneint werden. Vielmehr drängt sich das Bild von Benachteiligung auf mehreren Dimensionen auf. Die Kinder in diesem Verlaufstyp verfügen nicht nur über eine

unterdurchschnittliche Intelligenz, sondern auch über schwache Lernvoraussetzungen sowohl im sprachlichen als auch im mathematischen Bereich. Vielleicht wurden sie aus diesem Grund auch häufiger von der Einschulung zurückgestellt, was das durchschnittliche höhere Lebensalter in diesem Verlaufstyp nahelegen würde. Gleichzeitig sind diese Kinder interessanterweise aber auch sozial überdurchschnittlich integriert, was sich sowohl in den Beurteilungen ihrer Lehrpersonen zu ihrer sozialen Integration als auch in der häufigeren Teilnahme an strukturierten Freizeitaktivitäten widerspiegelt. Diese Aktivitäten stehen vielleicht mit den schulischen Aktivitäten zumindest am Anfang der Schulzeit noch in Konflikt und werden möglicherweise wegen der dort wahrscheinlicheren Erfolgserlebnisse von den Kindern auch favorisiert und auf Kosten von Schularbeiten priorisiert. Insgesamt ergibt sich also nicht das Bild des einen bestimmten Risikofaktors, sondern eher das eines Syndroms kombinierter Faktoren im Leistungs- und Sozialverhalten.

### **6.3 Theoretische und praktische Implikationen**

Die aufgezeigte Heterogenität und Komplexität der Verlaufsmuster von Lernzuwachsen macht deutlich, dass diese nur mit längsschnittlich angelegten Untersuchungen identifiziert und beschrieben werden können. Damit wird aber auch ganz deutlich die Begrenztheit von querschnittlichen Lernstandserhebungen aufgezeigt, weil hier beispielsweise zwei Schüler mit exakt gleichen Schulleistungen sich auf zwei völlig unterschiedlichen Verlaufspfaden befinden könnten, mit ganz unterschiedlichen Prognosen für den weiteren schulischen Erfolg. Die Identifikation unterschiedlicher Verlaufstypen und ihrer Korrelate kann daher hilfreich sein, um die Qualität der Prognose von Leistungsverläufen weiter zu verbessern, und zwar auch, wenn im Anwendungsfall nur eine einmalige Lernstandserhebung vorliegt. Vorausgesetzt, dass die identifizierten Verlaufsmuster auch über andere Bildungskontexte hinweg repliziert werden können, könnte man nämlich unter Zuhilfenahme der jetzt bekannten Typisierung den wahrscheinlichsten zukünftigen Verlauf bestimmen, wenn der aktuelle Lernstand und eine Reihe von zusätzlichen

Prädiktorvariablen bekannt sind. Allerdings sei an dieser Stelle angemerkt, dass es zum gegenwärtigen Zeitpunkt noch verfrüht erscheint, von allgemeingültigen Typen auszugehen, die sich in weiteren Studien finden lassen. Zukünftige Forschung muss zeigen, ob es über die Zeit und über Bildungssysteme hinweg vergleichbare Muster gibt oder nicht.

Theoretisch interessant bleibt aber trotzdem die Bestimmung, Erklärung und später auch die Modifikation der Prozesse, die der Manifestation der einzelnen Verlaufstypen zugrunde liegen. In den bisherigen Befunden deutet sich an, dass dafür das Angebots-Nutzungs-Modell von Helmke (2015) und verwandte Modelle hilfreich sein könnten. Für die Identifikation der wichtigsten Angebotsvariablen fehlen in dieser Untersuchung sowohl die Messungen von Schul- und Unterrichtsmerkmalen als auch eine ausreichende Heterogenität im Kontext (vgl. aber Helbling, Tomasik & Moser, in Druck). Für die Identifikation der wichtigsten Nutzungsvariablen deutet sich jedoch an, dass diese hauptsächlich mit den sprachlichen Kompetenzen der Schülerinnen und Schüler in Verbindung stehen könnten. Gerade in Bezug auf diesem Faktor wären Interventionsstudien denkbar, um der kausalen Struktur des Prozesses auf die Spur zu kommen. Auch der Bildungshintergrund der Eltern scheint eine Rolle zu spielen, wobei die Prozesse der primären und sekundären Herkunftseffekte schon recht gut verstanden sind (vgl. Maaz et al., 2009).

In praktischer Hinsicht ergibt sich neben der Anwendung für eine verbesserte Lernverlaufsprognose vor allen Dingen die Möglichkeit der personalisierten pädagogischen und/oder bildungspolitischen Förderung unter Berücksichtigung des zu erwartenden Verlaufstyps. So könnten beispielsweise Kinder, die ein hohes Risiko aufweisen, zum lernschwachen oder nachzügeln den Verlaufstyp zu gehören, besondere Unterstützung erhalten. Ebenso könnten die Angebote für die Kinder vom Verlaufstyp mit Vorsprung angepasst werden, damit sie stärker wie bisher von diesen Angeboten profitieren. Alle diese Punkte erfordern jedoch weitere Forschung zur besseren Prognostizierbarkeit der Verlaufstypenzugehörigkeit, möglichst früh in der schulischen Laufbahn der Kinder. Diese ist

nicht nur methodisch, sondern auch ethisch eine Herausforderung, wenn man vermeiden möchte, dass es zu frühen Festschreibungen und damit zu Reproduktionen von sozialen Ungleichheiten kommt, weil die Prognosen ja notwendigerweise auf früheren Daten beruhen, die das Gegebene, aber nicht unbedingt das Mögliche widerspiegeln.

### **6.3 Stärken und Schwächen der Studie**

Die vorliegende Studie verfügt über einige methodische Stärken, die sie im deutschsprachigen Raum bislang einzigartig machen. Betont werden soll neben der zufälligen Stichprobenziehung auf der Grundlage einer wohldefinierten Grundgesamtheit vor allen Dingen die Möglichkeit, den langfristigen Verlauf der Lernleistungen auf einer gemeinsamen metrischen Skala abzubilden. Die Verwendung standardisierter Leistungstests erlaubt es, die Schulleistungen unabhängig von Selbst- oder Fremdeinschätzungen zu messen, so dass Vergleiche auch über unterschiedliche Schultypen hinweg machbar sind. Außerdem sprechen der sehr geringe Stichprobenausfall und der Einbezug von Fremdeinschätzungen sowie Elterninformationen für die hohe Datenqualität. Auch stellt die verwendete statistische Methode der Mischverteilungsmodelle unseres Erachtens eine große Stärke der Studie dar, weil damit die große Heterogenität in den Schulleistungsverläufen adäquat berücksichtigt werden konnte.

Auf der anderen Seite lassen sich auch zumindest drei Schwächen der Studie identifizieren, die an dieser Stelle nicht unerwähnt bleiben sollen. Kritisiert werden kann zunächst sicherlich die enge inhaltliche Ausrichtung der verwendeten Leistungstests am Lehrplan des Kantons Zürich und damit an den in der Schule vermittelten Lerninhalten. Mit Absicht erfassen die gemessenen Schulleistungen nicht weitreichende Kompetenzen in den einzelnen Domänen, wie es etwa in den PISA-Studien der Fall ist. Damit wird im Zweifelsfall die Bedeutung von Prädiktoren überschätzt, die eng mit dem schulischen Lernen und Arbeiten in Zusammenhang stehen, wie etwa die Bildung der Eltern oder das selbständige Arbeitsverhalten der Kinder. Zweitens muss auf die mögliche Stichprobenabhängigkeit der

gefundenen Verlaufstypenlösung hingewiesen werden, zumal keine Möglichkeit besteht, diese an einer Replikationsstichprobe kreuzzuvalidieren. Hier ist tatsächlich zukünftige Forschung gefragt, die Lern- und Leistungsverläufe zu anderen Zeiten oder in anderen Kontexten zu untersuchen, um zu sehen, ob sich die postulierten Prototypen so oder so ähnlich wiederfinden lassen. Sollte das der Fall sein, dann wäre das ein starkes Indiz für die Gültigkeit der vorgestellten Lösung. Sollte das nicht der Fall sein, dann wäre es sehr spannend, die dafür verantwortlichen Kontextfaktoren zu identifizieren und zu verstehen. Drittens schließlich muss auf den korrelativen Charakter der Ergebnisse hingewiesen werden, auch wenn es fast selbstverständlich ist, dass in diesem Bereich experimentelle Studien sehr schwierig bis unmöglich durchzuführen sind. Der korrelative Charakter erlaubt aber keine Rückschlüsse in Bezug auf die Kausalität der untersuchten Faktoren zur Erklärung der Verlaufstypenzugehörigkeit. Hier könnten beispielsweise nicht untersuchte Drittvariablen kausal wesentlich relevanter sein und die gefundenen Korrelationen zum Teil vielleicht Scheinkorrelationen. Man kann versuchen dieser Frage in gut designten Interventionsstudien auf den Grund zu gehen, doch letztlich wird eine endgültige Klärung der Frage nach der Kausalität der Komplexität des untersuchten Sachverhalts wegen nur sehr schwer zu erreichen sein. Vielleicht ist das aber auch gar nicht nötig, wenn wir auch ohne dieses Wissen die Lernleistungen auf der Grundlage dieser Befunde positiv beeinflussen können.

#### **6.4 Ausblick**

Obwohl die Replikation unserer Befunde in anderen Stichproben oder Bildungskontexten noch aussteht, sollten bereits jetzt die Vorteile eines personenzentrierten Zugangs zu langfristigen Bildungsverläufen deutlich geworden sein. Zukünftige Forschung muss zeigen, wie fruchtbar dieser Zugang für Theoriebildung, Methodenentwicklung und Anwendung tatsächlich ist. Für die Theoriebildung erhoffen wir uns ein besseres Verständnis der Angebots- und Nutzungsvariablen, die zwischen den Verlaufstypen unterscheiden und innerhalb der Verlaufstypen mit Schulerfolg korreliert sind. Gerade auf den letzten Punkt sind

wir in unseren Analysen gar nicht eingegangen, obwohl sehr stark zu vermuten ist, dass für die einzelnen Verlaufstypen jeweils unterschiedliche Variablen und Prozesse für die Vorhersage des jeweiligen Schulerfolgs relevant sind. In der Methodenentwicklung wäre der nächste wichtige Schritt die Identifikation von Verlaufstypen und Typenzugehörigkeiten mit deutlich kürzeren längsschnittlichen Daten oder sogar einmaligen diagnostischen Instrumenten. Wie immer muss auch hier beachtet werden, dass sich die Struktur der Typologie und die relative Besetzung je nach Kontext verändern können, so dass auch hier mehr Forschung nötig ist. Und für die Anwendung, schließlich, wünschen wir uns ein besseres Verständnis dessen, wie sich die Typisierung auf individuelle Förderung auswirken kann. Gerade in einem gegliederten Schulsystem ist dabei die Frage berechtigt, ob sich mit einer angebots- und nutzungstheoretisch fundierten Typisierung nicht eine angemessenere und gerechtere Zuweisung zu den einzelnen Schulzweigen erreichen lässt.

#### Literatur

- Abt Gürber, N., Bucchieri, G., & Brühwiler, C. (2011). Engagement im Lesen und Lernstrategien. In Konsortium PISA.ch (Hrsg.), *PISA 2009: Regionale und kantonale Ergebnisse* (S. 25-36). Bern: BBT/EDK.
- Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin*, 102, 3-27. doi:10.1037/0033-2909.102.1.3
- Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, 77, 598-614. doi:10.1037/0021-9010.77.5.598
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317-332. doi:10.1007/BF02294359
- Angelone, D., Keller, F., & Moser, U. (2013). *Entwicklung schulischer Leistungen während der obligatorischen Schulzeit: Bericht zur vierten Zürcher Lernstandserhebung*

*zuhanden der Bildungsdirektion des Kantons Zürich*. Zürich: Bildungsdirektion Kanton Zürich.

- Angelone, D., & Moser, U. (2011). Die Zürcher Lernstandserhebung im Überblick. In Bildungsdirektion des Kantons Zürich (Hrsg.), *Nach sechs Jahren Primarschule: Deutsch, Mathematik und motivational-emotionales Befinden am Ende der 6. Klasse* (S. 119-128). Zürich: Bildungsdirektion Kanton Zürich.
- Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 329-341. doi:10.1080/10705511.2014.915181
- Baker, L., & Scher, D. (2002). Beginning reader's motivation for reading in relation to parental beliefs and home reading experiences. *Reading Psychology*, 23, 239-269. doi:10.1080/713775283
- Baumert, J., Gruehn, S., Heyn, S., Köller, O., & Schnabel, K.-U. (1997). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU): Dokumentation* (Bd. 1). Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J., Nagy, G., & Lehmann, R. (2012). Cumulative advantages of the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child Development*, 83, 1347-1367. doi:10.1111/j.1467-8624.2012.01779.x
- Baumert, J., & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323-407). Opladen: Leske + Budrich.
- Becherer, J., Köller, O., & Zimmermann, F. (2017). Sozialverhalten und Schulleistungen: Spielt die Beliebtheit in der Klasse eine Rolle? *Zeitschrift für Erziehungswissenschaft*,



20, 405-424.

- Bedard, K., & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, 121, 1437-1472. doi:10.1093/qje/121.4.1437
- Behörde für Schule und Berufsbildung (Hrsg.). (2011). *LAU: Aspekte der Lernausgangslange und der Lernentwicklung Klassenstufen 5, 7 und 9* (Hamburger Schriften zur Qualität im Bildungswesen Bd. 8). Münster: Waxmann.
- Behörde für Schule und Berufsbildung (Hrsg.). (2012). *LAU: Aspekte der Lernausgangslange und der Lernentwicklung Klassenstufen 11 und 13* (Hamburger Schriften zur Qualität im Bildungswesen Bd. 9). Münster: Waxmann.
- Bildungsdirektion Kanton Zürich. (2008). *Umsetzung Volksschulgesetz: Qualität in multikulturellen Schulen (QUIMS)*. Zürich: Lehrmittelverlag des Kantons Zürich. Abgerufen am 19.1.2018 unter [http://www.vsa.zh.ch/internet/bildungsdirektion/vsa/de/schulbetrieb\\_und\\_unterricht/qualitaet\\_multikulturelle\\_schulen\\_quims/\\_jcr\\_content/contentPar/downloadlist\\_1/downloaditems/751\\_1289910204702.spooler.download.1391004217588.pdf/quims\\_handreichung.pdf](http://www.vsa.zh.ch/internet/bildungsdirektion/vsa/de/schulbetrieb_und_unterricht/qualitaet_multikulturelle_schulen_quims/_jcr_content/contentPar/downloadlist_1/downloaditems/751_1289910204702.spooler.download.1391004217588.pdf/quims_handreichung.pdf).
- Bildungsdirektion Kanton Zürich. (2010). *Angebote für Schülerinnen und Schüler mit besonderem pädagogischen Bedürfnissen: Kleinklasse*. Zürich: Lehrmittelverlag des Kantons Zürich. Abgerufen am 19.1.2018 unter <http://www.vsa.zh.ch/dam/bildungsdirektion/vsa/schulbetrieb/sopaed/publikatione/Ordner%203/kleinklasse.pdf.spooler.download.1392989680849.pdf/kleinklasse.pdf>.
- Bloom, H. S., Hill, C. J., Black, A. R., & Lispey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, 1, 289-328. doi:10.1080/19345740802400072

- Blossfeld, H.-P., Bos, W., Hannover, B., Lenzen, D., Müller-Bölling, D., Prenzel, M., & Wößmann, L. (2009). *Geschlechterdifferenzen im Bildungssystem: Jahresgutachten 2009*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Blossfeld, H.-P., von Maurice, J., & Schneider, T. (2011). The national educational panel study: Need, main features, and research potential. *Zeitschrift für Erziehungswissenschaft*, 14, 5-17. doi:10.1007/s11618-011-0178-3
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York: Routledge.
- Bos, W., Lankes, E.-M., Schwippert, K., Valtin, R., Voss, A., Badel, I., & Plaßmeier, N. (Hrsg.). (2003). Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *Erste Ergebnisse aus IGLU: Schülerleistungen am Ende der vierten Jahrgangsstufe* (S. 69-141). Münster: Waxmann.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.
- Bourdieu, P. (1973). Kulturelle Reproduktion und soziale Reproduktion. In P. Bourdieu (Hrsg.), *Grundlagen einer Theorie der symbolischen Gewalt* (S. 88-137). Frankfurt am Main: Suhrkamp.
- Bourdieu, P., & Passeron, J.-C. (1971). *Die Illusion der Chancengleichheit: Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Stuttgart: Klett.
- Bradley, R. H., & Corwyn, R. F. (2006). The family environment. In C. Tamis-LaMonda & L. Balter (Eds.), *Child psychology: A handbook of contemporary issues* (2nd ed., pp. 493-520). New York: Garland.
- Brühwiler, C., Helmke, A., & Schrader, F.-W. (2017). Determinanten der Schulleistung. In M. K. W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktionen: Inhaltsfelder*,

- Forschungsperspektiven und methodische Zugänge* (S. 291-314). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Buckingham, D., & Willett, R. (Eds.). (2013). *Digital generations: Children, young people, and new media*. New York: Routledge.
- Budde, J. (2005). *Mathematikunterricht und Geschlecht: Empirische Ergebnisse und pädagogische Ansätze* (Bildungsforschung Band 30). Berlin: Bundesministerium für Bildung und Forschung.
- Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, 65, 23-35. doi:10.1007/s00265-010-1029-6
- Caprara, G. V., Barbanelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, 11, 302-306. doi:10.1111/1467-9280.00260
- Ceci, S. J., & Liker, J. K. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General*, 115, 255-266. doi:10.1037/0096-3445.115.3.255
- Chung, M. (2000). The development of self-regulated learning. *Asia Pacific Education Review*, 1, 55-66. doi:10.1016/j.lindif.2008.03.001
- Cobley, S., McKenna, J., Baker, J., & Wattie, N. (2009). How pervasive are relative age effects in secondary school education? *Journal of Educational Psychology*, 101, 520-528. doi:10.1037/a0013845
- Crawford, C., Dearden, L., & Greaves, E. (2013). *When you are born matters: Evidence for England* (IFS Report No. R80). London: Institute for Fiscal Studies. Abgerufen am 7.2.2018 unter <http://hdl.handle.net/10419/83505>.
- Cummins, J. (1991). Interdependency in first and second language proficiency in bilingual

- children. In E. Bialystock (Ed.), *Language processing in bilingual children* (pp. 70-89). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620652.006
- Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review*, 3, 101-129.  
doi:10.1016/j.edurev.2008.02.003
- Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1-26.  
doi:10.1016/j.edurev.2013.11.002
- Duncan, G. J., Magnuson, K. A., & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, 1, 59-80.  
doi:10.1080/15427609.2004.9683330
- Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, or marching band: What kind of extracurricular involvement matters? *Journal of Adolescent Research*, 14, 10-43. doi:10.1177/0743558499141003
- Eccles, J. S., & Templeton, J. (2002). Extracurricular and other after-school activities for youth. *Review of Research in Education*, 26, 113-180.  
doi:10.3102/0091732X026001113
- Elder, T. E., & Lubotsky, D. H. (2009). Kindergarten entrance age and children's achievement: Impact of state policies, family background, and peers. *The Journal of Human Resources*, 44, 641-683. doi:10.3368/jhr.44.3.641
- Ennemoser, M., Schiffer, K., Reinsch, C., & Schneider, W. (2003). Fernsehkonsum und die Entwicklung von Sprach- und Lesekompetenzen im frühen Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 35, 12-26.  
doi:10.1026//0049-8637.35.1.12

- Ferguson, D. M., Horwood, L. J., & Boden, J. M. (2008). The transmission of social inequality: Examination of the linkages between family socioeconomic status in childhood and educational achievement in young adulthood. *Research in Social Stratification and Mobility*, 26, 277-295. doi:10.1016/j.rssm.2008.05.001
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A five year longitudinal study. *Developmental Psychology*, 47, 1539-1552. doi:10.1037/a0025510
- Gentile, D. A., Swing, E. L., Lim, C. G., & Khoo, A. (2012). Video game playing, attention problems, and impulsiveness: Evidence of bidirectional causality. *Psychology of Popular Media Culture*, 1, 62-70. doi:10.1037/a0026969
- Gold, A., Duzy, A., Rauch, W. A., & Quiroga Murcia, C. (2012). Relatives Lebensalter und die Entwicklung schulischer Leistungen. *Zeitschrift für Bildungsforschung*, 2, 193-208. doi:10.1007/s35834-012-0046-0
- Gottfredson, L. S. (2002). G: Highly general and highly practical. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 525-538). Mahwah, NJ: Erlbaum.
- Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology*, 93, 3-13. doi:10.1037/0022-0663.93.1.3
- Gustafsson, J.-E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186-242). New York: Prentice Hall.
- Gut, J., Reimann, G., & Grob, A. (2012). Kognitive, sprachliche, mathematische und sozial-emotionale Kompetenzen als Prädiktoren späterer Schulleistungen: Können die Leistungen eines Kindes in den IDS dessen Schulleistungen drei Jahre später vorhersagen? *Zeitschrift für Pädagogische Psychologie*, 26, 212-220. doi:10.1024/1010-0652/a000070

- Helbling, L. A., Tomasik, M. J., & Moser, U. (in Druck). Long-term trajectories of academic performance in the context of social disparities: Longitudinal findings from Switzerland. *Journal of Educational Psychology*.
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (6. Aufl.). Seelze: Klett-Kallmeyer.
- Helsper, W., Kramer, R.-T., & Thiersch, S. (Hrsg.). (2014). *Schülerhabitus: Theoretische und empirische Analysen zum Bourdieuschen Theorem der kulturellen Passung*. Wiesbaden: Springer.
- Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, 111, 127-155. doi:10.1037/0033-2909.111.1.127
- Hopf, D. (2005). Zweisprachigkeit und Schulleistungen bei Migrantenkindern. *Zeitschrift für Pädagogik*, 51, 236-251.
- Jürges, H., & Schneider, K. (2006). Ins Frühjahr geborene Kinder haben schlechtere Bildungschancen. *Wochenbericht*, 17, 209-214.
- Keller, F., & Moser, U. (2008a). Die Untersuchung im Überblick. In U. Moser & J. Hollenweger (Hrsg.), *Drei Jahre danach: Lesen, Wortschatz, Mathematik und soziale Kompetenzen am Ende der dritten Klasse* (S. 13-40). Oberentfelden: Sauerländer.
- Keller, F., & Moser, U. (2008b). Fachleistungen am Ende der 3. Klasse. In U. Moser & J. Hollenweger (Hrsg.), *Drei Jahre danach: Lesen, Wortschatz, Mathematik und soziale Kompetenzen am Ende der dritten Klasse* (S. 41-87). Oberentfelden: Sauerländer.
- Klauer, K. J., & Leutner, D. (2007). *Lehren und Lernen: Einführung in die Instruktionspsychologie*. Weinheim: Beltz PVU.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices*. New York: Springer.
- Konsortium Bildungsberichterstattung. (Hrsg.). (2006). *Bildung in Deutschland: Ein*

- indikatorengestützter Bericht mit einer Analyse zur Bildung und Migration*. Bielefeld: Bertelsmann.
- Krohne, J. A., Meier, U., & Tillmann, K.-J. (2004). Sitzenbleiben, Geschlecht und Migration: Klassenwiederholungen im Spiegel der PISA-Daten. *Zeitschrift für Pädagogik*, 50, 373-391.
- Kuhn, H.-P. (2008). Geschlechterverhältnisse in der Schule: Sind die Jungen jetzt benachteiligt? Eine Sichtung empirischer Studien. In: B. Rendtoff & A. Prenzel (Hrsg.), *Kinder und ihr Geschlecht* (S. 49-71). Opladen: Verlag Barbara Budrich.
- Linacre, J. M. (1994). *Many-facet Rasch measurement*. Chicago: MESA Press.
- Lincove, J. A., & Painter, G. (2006). Does the age that children start kindergarten matter? Evidence of long-term educational and social outcomes. *Educational Evaluation and Policy Analysis*, 28, 153-179. doi:10.3102/01623737028002153
- Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88, 767-778. doi:10.1093/biomet/88.3.767
- Lücken, M., Thonke, F., Pohlmann, B., Hoffmann, B., Golecki, R., Rosendahl, J., ... Poerschke, J. (2014). KERMIT – Kompetenzen ermitteln. In D. Fickermann & N. Maritzen (Hrsg.), *Grundlagen für eine daten- und theoriegestützte Schulentwicklung: Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung* (Hamburger Schriften zur Qualität im Bildungswesen, Bd. 13, S. 127-154). Münster: Waxmann.
- Maaz, K., Baumert, J., & Trautwein, U. (2009). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? *Zeitschrift für Erziehungswissenschaft*, 12, 11-46. doi:10.1007/978-3-531-92216-4\_2
- Magnusson, D. (2003). The person approach: Concepts, measurement models, and research strategy. *New Directions for Child and Adolescent Development*, 101, 3-23.

doi:10.1002/cd.79

- Mahoney, J. L. (2000). School extracurricular activity participation as a moderator in the development of antisocial patterns. *Child Development, 71*, 502-516. doi:10.1111/1467-8624.00160
- Mahoney, J. L., Cairns, B. D., & Farmer, T. W. (2003). Promoting interpersonal competence and educational success through extracurricular activity participation. *Journal of Educational Psychology, 95*, 409-418. doi:10.1037/0022-0663.95.2.409
- Mahoney, J. L., & Stattin, H. (2000). Leisure activities and adolescent antisocial behavior: The role of structure and social context. *Journal of Adolescence, 23*, 113-127. doi:10.1006/jado.2000.0302
- Marsh, H. W. (1992). Extracurricular activities: Beneficial extension of the traditional curriculum or subversion of academic goals? *Journal of Educational Psychology, 84*, 553-562. doi:10.1037/0022-0663.84.4.553
- Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools, 43*, 267-281. doi:10.1002/pits.20149
- Matthews, G., Zeidner, M., & Roberts, R. D. (2006). Models of personality and affect for education: A review and synthesis. In P. A. Alexander & P. H. Winnie (Eds.), *Handbook of educational psychology* (2nd ed., pp. 163-186). Mahwah, NJ: Erlbaum. doi:10.1.1.570.2894
- McElvany, N., Becker, M., & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 41*, 121-131. doi:10.1026/0049-8637.41.3.121
- Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children and adolescents*. Mahwah, NJ: Lawrence Erlbaum Associates.



- Metsäpelto, R.-L., Pakarinen, E., Kiuru, N., Poikkeus, A.-M., Lerkkanen, M.-K., & Nurmi, J.-E. (2015). Developmental dynamics between children's externalizing problems, task-avoidant behavior, and academic performance in early school years: A 4-year follow-up. *Journal of Educational Psychology*, 107, 246-257. doi:10.1037/a0037389
- Morin, A. J. S., Rodriguez, D., Fallu, J.-S., Maïano, C., & Janosz, M. (2011). Academic achievement and smoking initiation in adolescence: A general growth mixture analysis. *Addiction*, 107, 819-828. doi:10.1111/j.1360-0443.2011.03725.x.
- Moser, U., & Angelone, D. (2011). Fachleistungen am Ende der 6. Klasse. In U. Moser, A. Buff, D. Angelone & J. Hollenweger (Hrsg.), *Nach sechs Jahren Primarschule: Deutsch, Mathematik und motivational-emotionales Befinden am Ende der 6. Klasse*. (S. 31-49). Zürich: Bildungsdirektion Kanton Zürich.
- Moser, U., Berweger, S., & Stamm, M. (2005a). Lesekompetenzen bei Schuleintritt. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz und soziale Kompetenzen beim Schuleintritt* (S. 37-58). Oberentfelden: Sauerländer.
- Moser, U., Berweger, S., & Stamm, M. (2005b). Mathematische Kompetenzen bei Schuleintritt. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz und soziale Kompetenzen beim Schuleintritt* (S. 77-98). Oberentfelden: Sauerländer.
- Moser, U., Berweger, S., & Stamm, M. (2005c). Wortschatz bei Schuleintritt. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz und soziale Kompetenzen beim Schuleintritt* (S. 59-76). Oberentfelden: Sauerländer.
- Moser, U., Berweger, S., & Tresch, S. (2003). *SMS: Sprache und Mathematik bei Schuleintritt* [Unveröffentlichter Test]. Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung.
- Moser, U., Oostlander, J., & Tomasik, M. J. (2017). Soziale Ungleichheiten im Leistungszuwachs und bei Bildungsübergängen. In M. P. Neuenschwander & C. Nägele

- (Hrsg.), *Bildungsverläufe von der Einschulung bis in den ersten Arbeitsmarkt: Theoretische Ansätze, empirische Befunde und Beispiele* (S. 59-77). Wiesbaden: Springer VS.
- Moser, U., & Stamm, M. (2005). Die Untersuchung im Überblick. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz, Mathematik und soziale Kompetenzen beim Schuleintritt* (S. 13-26). Oberentfelden: Sauerländer.
- Möble, T., Kleimann, M., & Rehbein, F. (2007). *Bildschirmmedien im Alltag von Kindern und Jugendlichen*. Baden-Baden: Nomos.
- Muller, P. A., Stage, F. K., & Kinzie, J. (2001). Science achievement growth trajectories: Understanding factors related to gender and racial-ethnic differences in precollege science achievement. *American Educational Research Journal*, 38, 981-1012.  
doi:10.3102/00028312038004981
- Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory class. *Alcoholism: Clinical and Experimental Research*, 24, 882-891. doi:10.1111/j.1530-0277.2000.tb02070.x
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus statistical analysis with latent variables: User's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- NICHD Early Child Care Research Network (2005). Predicting individual differences in attention, memory, and planning in first graders from experience at home, child care, and school. *Developmental Psychology*, 41, 99-114. doi:10.1037/0012-1649.41.1.99
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14, 535-569. doi:10.1080/10705510701575396
- OECD. (2015). *Was sind die Ursachen von Ungleichheit zwischen den Geschlechtern im Bildungsbereich?* (PISA im Fokus 2015/03). Paris: OECD. Abgerufen am 2.2.2018 unter [https://www.oecd.org/pisa/pisaproducts/pisainfocus/PIF-49%20\(ger\).pdf](https://www.oecd.org/pisa/pisaproducts/pisainfocus/PIF-49%20(ger).pdf).

- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36, 89-101. doi:10.1207/S15326985EP3602\_4
- Pekrun, R., vom Hofe, R., Blum, W., Götz, T., Wartha, S., Frenzel, A., & Jullien, S. (2006). Projekt zur Analysen der Leistungsentwicklung in Mathematik (PALMA): Entwicklungsverläufe, Schülervoraussetzungen und Kontextbedingungen von Mathematikleistungen in der Sekundarstufe I. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule: Abschlussbericht des DFG-Schwerpunktprogramms* (S. 21-53). Münster: Waxmann.
- Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research*, 84, 203-244. doi:10.3102/0034654313509492
- Puhani, P. A., & Weber, A. M. (2007). Does the early bird catch the worm? Instrumental variable estimates of educational effects of age of school entry in Germany. *Empirical Economics*, 32, 359-386. doi:10.1007/s00181-006-0089-y
- Ramaswamy, V., DeSarbo, W., Reibstein, S., & Robinson, W. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, 12, 103-124. doi:10.1287/mksc.12.1.103
- Ramm, G., Prenzel, M., Heidemeier, H., & Walter, O. (2004). Soziokulturelle Herkunft: Migration. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 254-272). Münster: Waxmann.
- Reiss, K., Klieme, E., Köller, O., & Stanat, P. (Hrsg.). (2017). PISA Plus 2012-2013: Kompetenzentwicklung im Verlauf eines Schuljahres [Sonderheft]. *Zeitschrift für Erziehungswissenschaft*, 20.
- Rhyn, H., Widmer, T., Roos, M., & Nideröst, B. (2002). *Zuständigkeiten und Ressourcen in Zürcher Volksschulen mit und ohne Teilautonomie (TaV): Evaluationsbericht*. Zürich:

- Kompetenzzentrum für Bildungsevaluation und Leistungsmessung. Abgerufen am 7.2.2018 unter <http://www.ibe.uzh.ch/dam/jcr:000000000-6ff9-ac1b-0000-000005e404ed/TaV.pdf>.
- Satin, A., & Shastry, W. (1990). *Survey sampling: A non-mathematical guide*. Ottawa: Minister of Supply and Services Canada.
- Schnabel, K.-U., & Schwippert, K. (2000). Einflüsse sozialer und ethnischer Herkunft beim Übergang in die Sekundarstufe II und den Beruf. In J. Baumert, W. Bos & R. H. Lehmann (Hrsg.), *TIMSS/III: Dritte internationale Mathematik und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn, Bd.1 Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit* (S. 261-300). Opladen: Leske + Budrich.
- Schneider, W., & Björklund, D. F. (1992). Expertise, aptitude, and strategic remembering. *Child Development*, 63, 461-473. doi:10.2307/1131492
- Schneider, W., Körkel, K., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology*, 81, 306-312. doi:10.1037/0022-0663.81.3.306
- Schröder-Lenzen, A., & Merken, H. (2006). Differenzen schriftsprachlicher Kompetenzentwicklung bei Kindern mit und ohne Migrationshintergrund. In A. Schröder-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung: Migration, Leistungsangst und Schulübergang* (S. 15-44). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- Sclove, L. S. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52, 333-343. doi:10.1007/BF02294360
- Spinath, B., Freudenthaler, H. H., & Neubauer, A. C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation.

- Personality and Individual Differences*, 48, 481-486. doi:10.1016/j.paid.2009.11.028
- Stecher, L. (2005). Informelles Lernen bei Kindern und Jugendlichen und die Reproduktion sozialer Ungleichheit. *Zeitschrift für Erziehungswissenschaft*, 8, 374-393. doi:10.1007/s11618-005-0146-x
- Stern, E. (1997). Mathematik. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Themenbereich D, Praxisgebiete, Serie I, Pädagogische Psychologie 3, S. 397-426). Göttingen: Hogrefe.
- Tomasik, M. J., Napolitano, C. M., & Moser, U. (in Druck). Trajectories of academic performance across compulsory schooling and thriving in young adulthood. *Child Development*. doi:10.1111/cdev.13150
- Totsika, V., & Sylva, K. (2004). The home observation for measurement of the environment revisited. *Child and Adolescent Mental Health*, 9, 25-35. doi:10.1046/j.1475-357X.2003.00073.x
- Tresch, S. (2005). Soziale Kompetenzen bei Schuleintritt. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz, Mathematik und soziale Kompetenzen beim Schuleintritt* (S. 99-112). Oberentfelden: Sauerländer.
- Tully, C. J. (1994). *Lernen in der Informationsgesellschaft: Informelle Bildung durch Computer und Medien*. Opladen: Westdeutscher Verlag.
- Waldmann, M. R., Renkl, A., & Gruber, H. (2003). Das Dreieck von Begabung, Wissen und Lernen. In W. Schneider & M. Knopf (Hrsg.), *Entwicklung, Lehren und Lernen: Zum Gedenken an Franz Emanuel Weinert* (S. 219-233). Göttingen: Hogrefe.
- Wang, M., & Bodner, T. E. (2007). Growth mixture modeling: Identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, 10, 635-656. doi:10.1177/1094428106289397
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54, 427-450. doi:10.1007/BF02294627

- Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, 43, 352-360. doi:10.3102/0013189X14553660
- Weber, H. S., Lu, L., Shi, J., & Spinath, F. M. (2013). The roles of cognitive and motivational predictors in explaining school achievement in elementary school. *Learning and Individual Differences*, 25, 85-92. doi:10.1016/j.lindif.2013.03.008
- Wei, X., Blackorby, J., & Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, 78, 89-106. doi:10.1177/001440291107800106
- Weinert, F. E. (1996). Wissen und Denken: Über die unterschätzte Bedeutung des Gedächtnisses für das menschliche Denken. In Bayerische Akademie der Wissenschaften (Hrsg.), *Jahrbuch 1996* (S. 85-101). München: Bayerische Akademie der Wissenschaften.
- Weinert, F. E., & Schneider, W. (1999). *Individual development from 3 to 12: Findings from the Munich longitudinal study*. Cambridge: Cambridge University Press.
- Weiß R., & Osterland, J. (1997). *Grundintelligenztest Skala I - CFT*. Göttingen, Germany: Hogrefe.
- Wentzel, K. R. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology*, 85, 357-364. doi:10.1037/0022-0663.85.2.357
- Winnie, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist*, 30, 173-187. doi:10.1207/s15326985ep3004\_2
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago: MESA Press.
- Zimmer, K., Burba, D., & Rost, J. (2004). Kompetenzen von Jungen und Mädchen. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G.

Rolff, J. Rost & U. Schiefle (Hrsg.), *PISA 2003: Der Bildungsstandard der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*.

Münster: Waxmann.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview.

*Educational Psychologist*, 25, 3-17. doi:10.1207/s15326985ep2501\_2

Zöllner, I., Ross, J., & Schöler, H. (2006). Einfluss soziokultureller Faktoren auf den

Schriftspracherwerb im Grundschulalter. In A. Schröder-Lenzen (Hrsg.),

*Risikofaktoren kindlicher Entwicklung: Migration, Leistungsangst und Schulübergang*

(S. 45-65). Wiesbaden: VS Verlag für Sozialwissenschaften.

*Tabelle 1.* Population und Stichprobe für die 1. Regelklassen sowie die Kleinklassen A

(ungewichtete Angaben)

	Anzahl in der Population		Anzahl in der Stichprobe	
	Klassen	Kinder	Klassen	Kinder
1. Regelklassen <sup>a</sup>	316	6226	40	813
1. Regelklassen „TaV“ <sup>b</sup>	159	3164	30	609
1. Regelklassen „QuimS“ <sup>c</sup>	30	562	30	516
1. Klassen der Kleinklassen	145	1166	20	105
Total	650	11118	120	2043

*Anmerkungen:* <sup>a</sup>ohne Mehrklassenschulen und Reformschulen <sup>b</sup>Modellprojekt „Teilautonome Volksschule“ <sup>c</sup>Modellprojekt „Qualität in multikulturellen Schulen“



*Tabelle 2.* Bildungssituation am Ende der obligatorischen Schulzeit neun Jahre nach Eintritt in eine 1. Regelklasse

	7. Klasse	8. Klasse	9. Klasse	10. Klasse	Total
Gymnasium		0.7%	17.3%	0.5%	18.5%
Abteilung A		4.5%	32.8%		37.2%
Abteilung B	0.1%	7.9%	24.3%		32.3%
Abteilung C	0.1%	1.9%	2.3%		4.2%
Kleinklasse/Sonderschule		1.2%		0.2%	1.3%
Privatschule	0.1%	1.6%	3.9%		5.5%
Berufslehre				0.3%	0.3%
Handels-/Fachmittelschule			0.6%	0.1%	0.7%
Total	0.2%	17.6%	81.1%	1.1%	100%

*Anmerkungen:* Das Sekundarschulsystem in Zürich war nach drei Abteilungen mit unterschiedlichem Anforderungsniveau gegliedert, wobei die Abteilung A das höchste Anforderungsniveau aufwies. Im Zuge von Schulreformen wurden die Abteilungen B und C dann später häufig zusammengelegt.

Tabelle 3. Deskriptive Kennwerte und Korrelationen der WLEs für Deutsch und Mathematik

Fach	<i>N</i>	<i>M</i>	Korrelationen (mit <i>SD</i> in der Diagonalen)					
			Deu 3	Deu 6	Deu 9	Mat 3	Mat 6	Mat 9
Deu 3	1859	497.28	102.73					
Deu 6	1644	780.11	.69	108.66				
Deu 9	1641	880.50	.63	.71	109.69			
Mat 3	1849	496.19	.65	.53	.51	102.46		
Mat 6	1650	873.48	.59	.68	.62	.64	111.74	
Mat 9	1642	926.05	.53	.62	.69	.61	.70	115.87

Tabelle 4. Deskriptive Verteilung der untersuchten Variablen ( $N = 2043$ )

	Mittelwert/ Anteil	Streuung	Minimum	Maximum
Geschlecht				
Jungen	50.8%			
Mädchen	49.2%			
Alter	6.95	0.37	5.65	8.60
Einschulung				
früh	3.0%			
normativ	87.6%			
spät	9.4%			
Intelligenz	100.36	14.90	70.00	145.00
Wortschatz	500	100	88.67	703.63
Lesen	500	100	299.03	705.31
Rechnen	500	100	166.09	853.31
Regelkonformität	3.15	0.60	1.00	4.00
Selbständigkeit	2.93	0.79	1.00	4.00
Kooperation	3.20	0.53	1.00	4.00
Integration	2.62	0.73	1.00	4.00
Fremdsprachigkeit	27.6%			
Bildungsdauer	12.58	3.52	4.00	17.00
Freizeitaktivitäten	0.00	1.00	-0.70	11.77
Familienzeit	0.00	1.00	-2.46	4.06
TV-Nutzung	1.67	0.91	0	4
Computernutzung	1.26	0.97	0	3

*Tabelle 5.* Deskriptive Beschreibung der effektiven Stichprobe zu den vier Messzeitpunkten.

	Schuljahr	Schuljahr	Schuljahr	Schuljahr
	2003/04	2006/07	2009/10	2012/13
	(T1)	(T2)	(T3)	(T4)
Anzahl	2043	1959	1803	1634
Anteil Mädchen	49.2%	49.7%	50.1%	50.3%
Alter zur Einschulung <sup>a</sup>	6.95	6.94	6.95	6.94
Anteil fremdsprachiger Kinder	27.6%	27.7%	27.7%	28.2%
Höchste Bildungsdauer der Eltern <sup>a</sup>	12.58	12.60	12.59	12.55

*Anmerkungen:* <sup>a</sup> Angaben in Jahren.

Tabelle 6. Güte der Modellpassung für unterschiedliche Anzahl von Typen

$k$	LL	LL( $k-1$ ) <sup>a</sup>	2 $\Delta$ LL <sup>a</sup>	$p^a$	AIC	BIC	SSABIC	Entropie
1	10250.46				20534.917	20628.05	20574.04	
2	10206.74	10986.36	1559.25	.003	20457.473	20577.99	20508.10	.532
3	10182.16	10939.50	1514.68	.451	20418.314	20566.23	20480.45	.531
4	10161.87	10913.15	1502.56	.527	20387.747	20563.05	20461.39	.514
5	10149.09	10891.42	1484.65	.001	20372.178	20574.87	20457.32	.528
6	10142.57	10877.71	1470.28	.627	20369.148	20599.23	20465.80	.525

*Anmerkungen:* <sup>a</sup>Berechnung der Log-Likelihood-Differenz nach Lo, Mendell und Rubin

(2001), weil die Differenz der gewöhnlichen Loglikelihoods (z. B. 10250.46 – 10206.74)

nicht  $\chi^2$ -verteilt ist und deswegen eine „korrigierte“ Loglikelihood für die Berechnung der

Differenz verwendet werden muss (z. B. 10986.36 – 10206.74).

Tabelle 7. Anteile und Wachstumsparameter der fünf identifizierten Verlaufstypen

Typenbezeichnung	Anteil	Schulleistungen		Mathematik- Residuum	
		Intercept	Slope	Intercept	Slope
Nachzügler	8.9%	436 (17)	106 (5)	-26 (16)	27 (5)
Lernschwache	26.3%	451 (20)	116 (3)	-37 (9)	30 (3)
Durchschnitt	20.1%	472 (20)	124 (3)	59 (13)	19 (4)
Ertragsreich	30.1%	512 (12)	142 (3)	-55 (8)	49 (1)
Vorsprung	14.6%	617 (14)	122 (4)	62 (14)	20 (4)
ohne Mischverteilungsmodell	100%	497 (6)	125 (1)	-1 (4)	31 (1)

*Anmerkungen:* Standardfehler in Klammern; Lesebeispiel: Ohne Mischverteilungsmodell

betragen die gesamten Schulleistungen 497 Punkte und steigen (in der Grundschule, später findet eine Abflachung des Zuwachses statt) um 125 Punkte pro Schuljahr an. Die Schulleistungen in Deutsch und Mathematik unterscheiden sich anfangs in ihren Punktwerten nicht. Der Lernzuwachs in Mathematik ist aber jährlich um 31 Punkte grösser als der in Deutsch, so dass die Leistungen in Mathematik (in der Grundschule, später findet eine Abflachung des Zuwachses statt) um  $125 + 31 = 156$  Punkte jährlich steigen. Der Anteil der Schülerinnen und Schüler mit Startvorteil (um ein Beispiel im Mischverteilungsmodell herauszugreifen) beträgt 15%, ihre gesamten Schulleistungen sind anfangs mit 617 Punkten sehr hoch und der Anstieg bewegt sich mit 122 Punkten im durchschnittlichen Bereich. Relativ zu den gesamten Schulleistungen sind ihre Mathematikleistungen anfangs mit 62 zusätzlichen Punkten besonders hoch, steigen aber über den gesamten Anstieg hinaus etwas langsamer an als im Durchschnitt.

*Tabelle 8. Zusammenhang mit der Wahrscheinlichkeit der Verlaufstypenzugehörigkeit und Variablen auf der Schülerebene*

	Nachzügler	Lernschwach	Durchschnitt	Ertragsreich	Vorsprung
Geschlecht <sup>a</sup>	-.10	.15	-.20	.21	-.14
Alter	.13			-.16	
Früheinschulung <sup>b</sup>		-.10			
Späteinschulung <sup>b</sup>			.10	-.11	
Intelligenz	-.36	-.30	-.13	.24	.45
Wortschatz	-.33	-.16	-.21	.28	.30
Lesen	-.26	-.17	-.26	.29	.30
Rechnen	-.28	-.28	-.13	.18	.44
Regelkonformität	-.19		-.13	.17	.15
Selbstständigkeit	-.25	-.23	-.16	.25	.30
Kooperation	-.16		-.13	.13	.19
Integration		-.14		.11	.12

*Anmerkungen:* Dargestellt sind nur Korrelationen mit  $|r| \geq .10$ ; <sup>a</sup> Referenzkategorie

„männlich“; <sup>b</sup> Referenzkategorie „normative Einschulung“

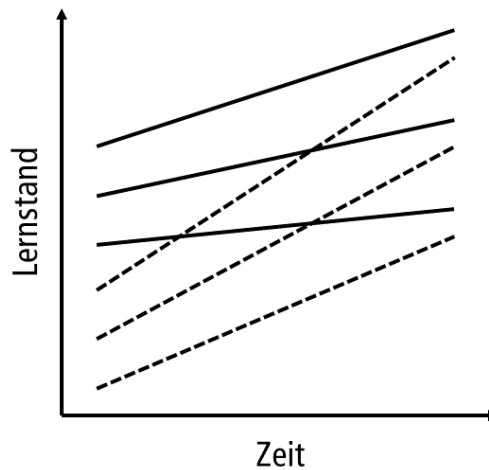
*Tabelle 9.* Zusammenhang mit der Wahrscheinlichkeit der Verlaufstypenzugehörigkeit und Variablen auf der Ebene des Herkunftskontexts

	Nachzügler	Lernschwach	Durchschnitt	Ertragsreich	Vorsprung
Fremdsprachigkeit	.24	.10	.13	-.21	-.17
Bildungsdauer	-.25	-.21	-.16	.27	.25
Freizeitaktivitäten	.13				
Familienzeit	-.12	-.10		.16	
TV-Nutzung	.13	.15	.12	-.18	-.15
Computernutzung		-.10			.14

*Anmerkungen:* Dargestellt sind nur Korrelationen mit  $|r| \geq .10$ .

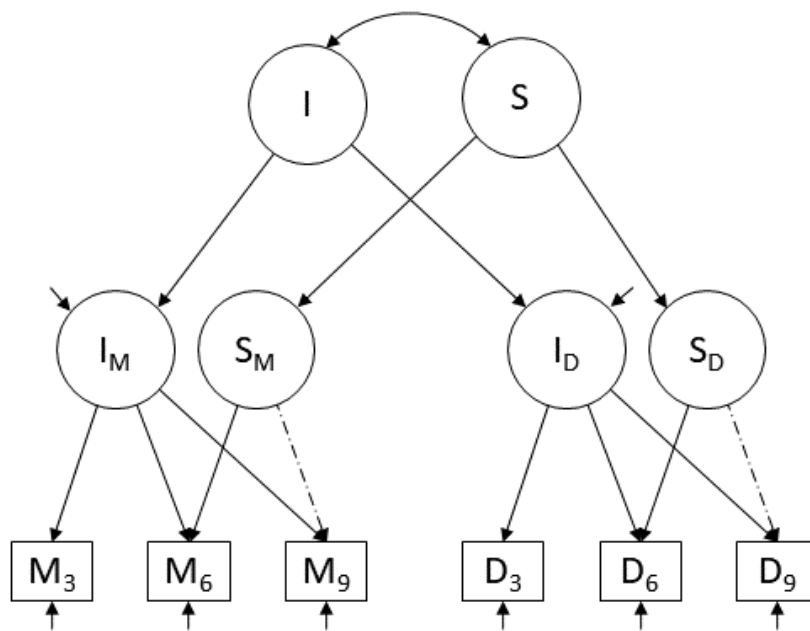


Abbildung 1. Hypothetisches Beispiel für unterschiedliche Schlussfolgerungen bei gesamthafter und typenspezifischer Betrachtung



*Anmerkungen:* In diesem hypothetischen Beispiel existiert für die Gesamtgruppe eine negative Korrelation zwischen Ausgangsniveau und Lernzuwachs. Lernende mit einem höheren Ausgangsniveau zeigen einen langsameren Lernzuwachs, die Ungleichheiten verkleinern sich in der Gesamtgruppe. Innerhalb der zwei unterschiedlichen Typen (dargestellt durch unterschiedliche Schraffierung) gilt genau der umgekehrte Zusammenhang, so dass die Korrelation positiv ist. Lernende mit höherem Ausgangsniveau zeigen einen höheren Lernzuwachs, die Ungleichheiten vergrößern sich innerhalb der anfänglich schwächeren bzw. anfänglich stärkeren Schüler.

Abbildung 2. Verwendetes latentes Wachstumskurvenmodell



*Anmerkungen:* Intercept (I) und Slope (S) für die Fächer Mathematik (M) und Deutsch (D), wobei die Koeffizienten der durchgezogenen Pfeile auf 1 festgesetzt wurden und die Koeffizienten der gestrichelten Pfeile frei aus den Daten geschätzt wurden.